Society for the Study
of Economic Inequality

Working Paper Series

Dividend tax credits and the elasticity of taxable income: evidence from small businesses

Pablo Gutierrez Cubillos

ECINEQ 2022 630

ECINEQ 2022 630
2022 December
www.ecineq.org

Dividend tax credits and the elasticity of taxable income: evidence from small businesses

Pablo Gutierrez Cubillos
University of Chile

Abstract

We assess firms' taxable income response to a dividend tax credit increase whencorporate and personal taxes are integrated. First, we theoretically show that, in anintegrated tax system, welfare changes stemming from a rise in corporate taxes dependon two parameters: the elasticity of taxable income with respect to the corporate taxrate and with respect to the dividend tax credit. Second, to estimate both parameters,we propose an identification strategy that relies on the bunching methodology and theexcess bunching difference before and after a tax reform that increased the dividendtax credit. Using Canadian administrative tax data and the presence of a kink inthe corporate tax system, we estimate these elasticities and empirically show that theincrease in the dividend tax credit reduced the deadweight loss associated with anincrease in the corporate tax by more than 50%. Our results are robust to a battery ofrobustness checks, including nonparametric estimates of the counterfactual densityin the bunching procedure.

Keyword: Bunching Estimation, Corporate taxation, Dividend taxation, Elasticity of taxable income, Small business, Tax integration

JEL Cassification: H25

Dividend tax credits and the elasticity of taxable income: evidence from small businesses

Pablo Gutiérrez Cubillos*

December 15, 2022

Abstract

We assess firms' taxable income response to a dividend tax credit increase when corporate and personal taxes are integrated. First, we theoretically show that, in an integrated tax system, welfare changes stemming from a rise in corporate taxes depend on two parameters: the elasticity of taxable income with respect to the corporate tax rate and with respect to the dividend tax credit. Second, to estimate both parameters, we propose an identification strategy that relies on the bunching methodology and the excess bunching difference before and after a tax reform that increased the dividend tax credit. Using Canadian administrative tax data and the presence of a kink in the corporate tax system, we estimate these elasticities and empirically show that the increase in the dividend tax credit reduced the deadweight loss associated with an increase in the corporate tax by more than 50%. Our results are robust to a battery of robustness checks, including nonparametric estimates of the counterfactual density in the bunching procedure.

Keywords: Bunching Estimation, Corporate taxation, Dividend taxation, Elasticity of taxable income, Small business, Tax integration.

JEL Classification: H25.

*Vancouver School of Economics, University of British Columbia. Email: pgutiecu@mail.ubc.ca

1 Introduction

Governments depend on different tax instruments to collect revenues, such as personal, corporate, sales, and capital gain taxes. This is because i) there are distributive consequences of choosing a particular tax base, and ii) there are behavioural effects associated with an increase in a tax rate. In addition, there are tax credits or benefits that decrease the tax base and incentivize agents to change their behaviour. For example, the accelerated depreciation tax benefit tries to increase firms' investment. Another relevant example is that, to avoid double taxation, some countries integrate the corporate and the dividend tax by giving a dividend tax credit that depends (sometimes notionally) on the amount of corporate tax paid.[1] This context raises the question of whether this type of integration produces behavioural effects in addition to the mechanical effect of decreasing double taxation.

To answer this question, we can estimate the effect of the dividend tax credit on the taxable income declared by firms. As Feldstein (1995, 1999) argue, the elasticity of taxable income (ETI) may be a sufficient statistic to measure the effect on welfare after a tax change. Intuitively, the magnitude and the sign of this elasticity shed light on firms' response to a credit that is paid outside the firm.[2] Thus, our first contribution to the literature is to show that in an integrated tax system in which the corporate tax rate and the dividend tax rate are integrated, the elasticity of taxable income with respect to the corporate tax rate is not a sufficient statistic to compute welfare changes after an increase in the corporate tax rate. An additional necessary statistic is the elasticity of taxable income with respect to the dividend tax credit. Intuitively, an increase in the corporate tax rate incentivizes firms to reduce their taxable income. In contrast, the increase in the dividend tax credit after an increase in the corporate tax rate may generate incentives to pay more dividends, hence increasing the taxable income declared by firms. Thus, the net effect of increasing the corporate tax should account for both effects. We theoretically show this by extending and complementing the framework of Devereux et al. (2014) when there is tax integration. This contribution aligns with Slemrod and Kopczuk (2002), who states that the government can indirectly choose an "optimal" ETI by changing tax parameters. In our context, the government can modify the effect of the corporate tax rate on welfare by

[1] For example, some OECD countries that have some integration between the corporate and the personal tax rate include Australia, Canada, Chile, Colombia, Korea, Mexico, New Zealand, and the UK.

[2] Firms' response to tax changes at the personal level are related to what the literature calls "the corporate veil".

changing the dividend tax credit. Moreover, our contribution follows Saez et al. (2012) in the sense that the ETI can be insufficient as the only parameter to identify the deadweight loss due to a tax change. Additionally, our contribution is based on Waseem (2018), who estimates the welfare effect of a tax change on other tax bases. In our case, we estimate welfare changes that follow from the increase of dividend tax credit after a rise in the corporate tax.

To estimate this elasticity, we use the Canadian tax system. In particular, small firms are taxed under a "kink" in the corporate tax rate — i.e. a change in the marginal tax rate. Below a certain threshold, firms pay a reduced corporate tax rate called the small business deduction (SBD), while above the threshold, firms pay the general corporate tax rate, which is higher than the SBD. Dividends paid with taxable income taxed under the SBD are fully integrated with the corporate tax. This means that all the corporate tax paid by those firms is a tax credit against the dividend tax paid by the firms' owners. However, dividends taxed at the general corporate tax rate are only partially integrated. That is, firm owners whose firms paid the general corporate tax rate only received a partial refund for the corporate tax paid by the firm.[3] In 2005, the Canadian parliament introduced a tax reform that increased the dividend tax credit received by firm owners whose firms are taxed under the general tax rate. This meant an increase in the integration rate for firms that paid the general corporate tax rate. This reform was implemented in 2006.

In this context, our second contribution is to extend the bunching estimation framework by showing that, under plausible assumptions which are conditional on the Canadian tax system and the 2006 tax reform, the elasticity of taxable income with respect to the dividend tax credit is identified and can be estimated through a difference in bunching approach. Intuitively, it is possible to use the difference in the relative excess bunching before and after the 2006 dividend tax reform to estimate the parameter of interest. This contribution extends the empirical framework from Devereux et al. (2014) that is based on Saez (2010) and Chetty et al. (2011).

Our third contribution is that, to the best of our knowledge, this is the first work to estimate the elasticity of taxable income with respect to the dividend tax credit. We do this for Canada using administrative tax data at the business level. We estimate that the elasticity of one minus the dividend tax credit is between -0.4 and -0.5, which is lower in absolute value than the elasticity of one minus the corporate tax rate, which ranges

[3] The dividend tax credit was close to the SBD corporate tax rate independent of the actual corporate tax paid by the firm.

from 0.6 to 0.7. This implies that firms react more strongly to corporate tax changes than dividend tax credit changes. Thus, we extend the corporate ETI literature (Bachas and Soto, 2021; Boonzaaier et al., 2019; Coles et al., 2021; Devereux et al., 2014; Lediga et al., 2019; Lobel et al., 2021) and the Canadian dividend tax credit literature (Smart, 2017, 2021). With these estimates, we compute that the behavioural effect of a one percentage point increase in the corporate tax is -38% of the mechanical effect. Additionally, we find that following the 2006 tax reform, this effect falls to -15%. That is, the increase in the dividend tax credit decreases the deadweight loss associated with an increase in the corporate tax. These results are robust to different counterfactual distribution assumptions, including a non-parametric estimate for the counterfactual distribution following the critique by Blomquist et al. (2021).

This paper proceeds as follows: Section 1.1 gives a literature review of the ETI and the bunching estimator. Section 2 describes the institutional background. Section 3 describes the theoretical framework. Section 4 presents the empirical strategy as well as the results. Section 5 is a policy discussion and Section 6 concludes.

1.1 Related literature

1.1.1 The relevance of the elasticity of taxable income

The elasticity of taxable income is a key parameter to estimate the deadweight loss associated with changes in taxes. As discussed by Feldstein (1999), the compensated elasticity of taxable income summarizes information regarding tax avoidance that reflects a richer picture of the effect of taxes on welfare.[4] Slemrod and Kopczuk (2002) further study the determinants of the elasticity of taxable income. They argue that the elasticity of taxable income is a function of preferences and the other margins of response to taxation.[5] Their main contribution is that there is an optimal level of the elasticity of taxable income that the government's tax rules can implicitly select. Using a theoretical framework, they identify three types of effects that any policy instrument has: a direct welfare cost, the revenue effect holding the strength of behavioural response constant, and

[4]Indeed, Feldstein (1999) argues that the previous deadweight loss analysis done by Harberger (1964) was incomplete because it uses the labour supply elasticity instead of the elasticity of taxable income. The latter incorporates information on tax avoidance through changes in compensations like employment pay health insurances and on changes of patterns of consumption.

[5]For example, the enforcement regime, the ability to change the business organization, and the capital gains realizations depend on the tax code.

the impact on distortions caused by other tax instruments. Given the benefits and the cost of implementing tax rules, there is an optimal elasticity of the taxable income.

This point was further explored and explained by Saez et al. (2012). They study three cases in which the elasticity of taxable income is an insufficient parameter to identify the deadweight loss associated with tax changes. First, when decreases in tax revenues are offset by increases in the future or other tax bases.[6] Second, when there are externalities associated with the responses to taxation.[7] Third, when the elasticity depends on the tax system. That is, a high ETI does not necessarily imply that there should be a reduction in the tax rate. For instance, in tax systems with narrow bases and many deductions or avoidance opportunities, it is preferable to broaden the tax base and eliminate avoidance opportunities to reduce the ETI than to lower tax rates.

Indeed, the literature has recently focused on adding more than just the ETI to account for behavioural responses to tax rate changes. For instance, Piketty et al. (2014) include tax avoidance effects as a part of the ETI of high-income earners. They distinguish between different behavioural responses such as tax avoidance, real responses and bargaining channels. le Maire and Schjerning (2013) and Kreiner et al. (2014, 2016) show that intertemporal income-shifting explains why the short-run ETI is low among self-employed and top income earners in Denmark. This implies that changes in real economic activity might be small even though the estimated ETI is considerable high. Best et al. (2015) study the effects of alternative tax policies such as broadening the base. They show that this policy distorts production proportionally to the elasticity of real output with the tax rate. Additionally, it reduces the returns to tax evasion proportionally to the evasion elasticity.

Devereux et al. (2014) show that income-shifting between corporate and personal tax bases plays a significant role in explaining the behavioural responses from corporate tax changes, while Harju and Matikka (2014) show that, absent any real effects, income-shifting between tax bases is active among the main owners of privately held corporations in Finland. Bachas and Soto (2021) find a large ETI for corporate income taxes (in the order of 3 to 5). They decompose the ETI between revenue and expenditure elasticities, showing that the tax avoidance behaviour that stems from the rise in the corporate tax rate is explained by an increase in expenditures. Boonzaaier et al. (2019) and Lediga et al.

[6]In particular, Chetty (2009) shows that in the case of income sheltering, when there is significant scope for income shifting, the ETI of an income source will not be a sufficient statistic to perform welfare analysis. Specifically, shifted income may generate offsetting increases in tax revenue collected from other income streams.

[7]Saez et al. (2012) illustrate this by using charitable giving behaviour as an example.

(2019) estimate the elasticity of taxable income with respect to the corporate tax for South Africa. These papers show that tax avoidance behaviour related to increased corporate taxes is driven by underreporting of sales rather than by over expenditure. Waseem (2018) develops an analytic framework in which the welfare responses are driven by real responses, evasion responses and income shifting responses. He estimates an ETI of the corporate taxable income of 2 for Pakistan. In addition, he shows that after accounting for the fiscal externalities of this reform, the welfare cost due to the tax reform in Pakistan increased by 40%.

1.1.2 The bunching estimator

We use the bunching methodology to estimate the ETI, a method that was first used by Saez (1999) and then by Saez (2010). It consists of using behavioural manipulation below tax kinks, which are discontinuities in the marginal tax rate, to estimate the excess mass of people located close to this kink. With this estimate, and under the counterfactual assumption that, without tax kinks, taxpayers would be uniformly distributed close to the kink, it is possible to estimate the ETI. This methodology was further extended by Chetty et al. (2011), who used a polynomial counterfactual and adjusted the excess mass to the right of the counterfactual distribution. Later, Kleven and Waseem (2013) extended the methodology for notches, which are changes in the average tax rate.

The developments of this literature are summarized by Kleven (2016), who argues that there are two broad lessons from the existing bunching literature. First, moving from observed bunching to a structural parameter that allows for predicting the effects of policy changes is difficult due to a range of optimization frictions that attenuate bunching and are challenging to observe and model. For example, hours constraints, search costs, adjustment costs, inattention, and uncertainty. Indeed, any evidence of sharp bunching in earnings likely results from tax evasion or tax avoidance rather than real responses.Second, the caveats when estimating structural elasticities do not invalidate the bunching approach. Still, they imply that the canonical approach might be used to study outcomes that are less subject to optimization friction. In this context, Gelber et al. (2020) developed a procedure to estimate adjustment costs and the elasticity of taxable income. They applied it to the Security Earnings Test, showing that changing a tax rate can be lower than was previously estimated. In addition, He et al. (2021), using data from China, finds that the ETI estimates from the bunching method are much lower than the long-run estimates from the quasi-experimental tax reform approach.

Moreover, the identification of the bunching approach has been questioned. For instance, Blomquist et al. (2021) show that the bunching estimator cannot identify the ETI without additional assumptions, as there is one equation for two unknowns: i) the ETI and ii) the preferences. They show that, for any excess mass around the bunching windows, there is a preference that supports any value for the ETI. Concretely, using a uniform distribution or a polynomial to estimate the counterfactual distribution implicitly defines the preferences and thus the ETI. Hence, an estimation procedure under those assumptions cannot be named a non-parametric procedure. This point is also made by Bertanha et al. (2021).[8] In this regard, there are multiple ways to account for this issue, one is to use non-parametric bounds for the ETI, and the other is to use a different group of non-bunchers as a counterfactual. For instance, this could be a different year where the threshold has changed (Devereux et al., 2014) or a different age group where there is no incentive for earnings manipulation at the threshold (Gelber et al., 2020).

Regarding the ETI in the context of the corporate tax rate, there are estimates for the UK (Devereux et al., 2014), the US (Coles et al., 2021), Costa Rica (Bachas and Soto, 2021), Honduras (Lobel et al., 2021), and South Africa (Boonzaaier et al., 2019; Lediga et al., 2019). These estimates range from 0.13 to 4. depending on the point where the elasticity is computed and how enforceable the corporate tax is. This elasticity is generally lower for developed countries than for developing countries.

2 Institutional background

2.1 Canadian corporate tax system

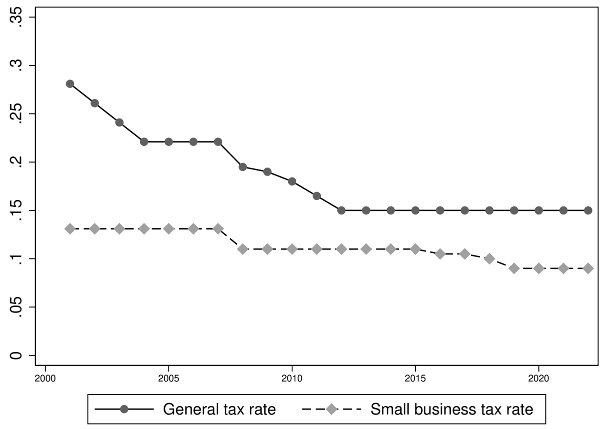
In Canada, firms pay corporate taxes depending on their structure, assets and location. Publicly traded and foreign-controlled firms pay the general business tax rate on all of their profits. In contrast, Canadian controlled private corporations (CCPCs) with less than $15 million[9] in taxable capital pay the substantially lower "small business tax rate" on profits up to the prescribed federal threshold — profits in excess of this threshold are

[8] They state that the point identification is impossible if the incentive schedule only contains kinks because there always exists an unobserved distribution that reconciles any elasticity with the observed distribution of responses. This is the same argument made by Blomquist et al. (2021).

[9] The $ symbol denotes Canadian dollars (CAD) unless stated otherwise.

taxed at the general rate.[10] This discounted tax rate is called the small business deduction (SBD)[11]. Figure 1 shows the historical trends of the small business tax rate and the general tax rate.

Figure 1: Small business tax rate and general tax rate†.



Source†: Finances of the Nation tax database

We can see that both tax rates have decreased between 2000 and 2022. In addition, there was a substantial decrease in the difference between rates between 2000 and 2015. However, since 2015 this difference has steadily increased. The small business tax rate is applicable until a threshold on the active income of firms.[12] We will call this threshold the

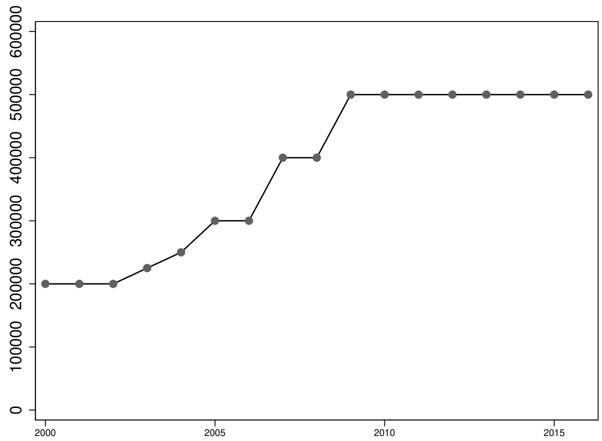
[10] A firm is a Canadian Controlled Private Corporation if the following conditions met: i) it is a private corporation, ii) it is a corporation that resides in Canada and was either incorporated in Canada or resident in Canada from June 18, 1971, to the end of the tax year, iii) it is not controlled directly or indirectly by one or more non-resident persons, iv) it is not controlled directly or indirectly by one or more public corporations (other than a prescribed venture capital corporation, as defined in Regulation 6700 of the Income Tax Regulations), v) it is not controlled by a Canadian resident corporation that lists its shares on a designated stock exchange outside of Canada vi) it is not controlled directly or indirectly by any combination of persons described in the three previous conditions, vii) if all of its shares that are owned by a non-resident person, by a public corporation (other than a prescribed venture capital corporation), or by a corporation with a class of shares listed on a designated stock exchange were owned by one person, that person would not own sufficient shares to control the corporation, viii) no class of its shares of capital stock is listed on a designated stock exchange.

[11] The SBD is applicable for firms whose taxable capital is lower than $15 million, but it is reduced in a straight line base when the taxable capital higher than $10 million. In addition, for a legal definition of taxable capital, refer to Income Tax Act (R.S.C., 1985, c. 1 (5th Supp.))

[12] Active income is a concept closely related to taxable income. Generally, active business income is income earned from a business source, including any income incidental to the business. Income from a specified investment business or a personal services business and income described in subparagraph (a)(i) of the definition of specified corporate income in subsection 125(7) for the year are generally not considered active business income and are not eligible for the SBD.

small business deduction threshold. Figure 2 shows the value of this threshold between 2000 and 2022. In this threshold there is a kink, which implies that after the threshold, there is a change on the marginal corporate tax rate that firms pay when their taxable income is higher than the SBD threshold.

Figure 2: Federal corporate threshold for small business deduction in $.†



Source†: Canadian tax simulator, Milligan (2016).

Figure 3 shows the kink in the corporate tax rate in 2006 when the SBD threshold was $300,000. For instance, a CCPC that declares $200,000 in active income pays $26,260 in taxes as only the small business tax rate (the corporate tax rate after applying the small business deduction) is applied to its active income. In contrast, a CCPC entitled to the small business deduction that declares $400,000 in active income firm must pay $39,390 + $22,130 (i.e., $300,000×13.13% and $100,000×22.13%).

Figure 3: Corporate tax rate scheme for CCPCs that are entitled to the small business deduction for 2006.

This kink generates incentives for tax planning since some firms locate close to the left of the threshold to avoid the payment of the general tax rate. Theoretically, we should observe bunching close to the federal threshold.

2.2 Canadian dividend tax system.

Since 1949 the Canadian tax system has been partially integrated, meaning that a dividend tax credit is given to shareholders when they receive their post corporate tax dividends from the firm. Under an integrated tax system, dividends are taxed at a lower tax rate to make the combined corporate and dividend tax rate close to the personal tax rate on other forms of income.

Since 1972, Canadian integration has been achieved through a "gross-up-and-credit" system that aims to reduce double taxation while fully taxing the underlying corporate income at each shareholder's own personal marginal tax rate. Integration thus works in two steps: (i) From the after corporate tax dividend amount (div_t), a gross-up rate (c) is applied to the dividend. (ii) A dividend tax credit is applied to the grossed-up amount. The final amount is the dividend tax credit, and the total tax paid is the personal tax rate multiplied by the grossed up level of dividends minus the DTC. Before 2006, the Canadian dividend tax system was integrated only for the small business tax rate, leading to incomplete integration for firms that paid the general corporate tax rate.[13] To illustrate this, Table 1 shows an example for small and large corporations.

[13]This applies to non-CCPCs and CCPCs that generate taxable income above the federal threshold.

Table 1: Dividend tax system prior to 2006 reform. Active income equal to 1,000.

	General	Small business
Federal corporate tax (-)	221	131
Dividend (net of corporate tax receive)	779	869
Gross up dividend (+)	896	1000
Personal tax paid (-)	260	290
DTC (+)	131	131
Dividend tax paid	129	159
Dividend received	650	710

Assume that a firm owned by one shareholder generates active income equal to $1000. This firm faces the general corporate tax rate. First, the firm pays $221 as corporate tax (22.1% of corporate tax rate). Then, the shareholder receives $779 as dividends net of the corporate tax rate. A gross-up rate is applied to those dividends, transforming them into $896. To this amount, the personal tax is charged (29% rate), which gives an amount of $260 ($896*0.29). Finally, a dividend tax credit equal to 13.1% of the taxable income from the firm is applied to the $260. Thus, the net dividend tax paid equals $129 ($260-$131), and the dividend net of taxes received by the shareholder is equal to $650 - dividends net of corporate tax ($779) - minus personal tax paid on dividends net of dividend tax credit ($129). However, there is some residual double taxation because the tax paid by the shareholder in the case of receiving the taxable income generated by the firm ($1000) is equal to $290, which is lower than the total tax paid between the firm and the shareholder ($350=$221+$129). As we can see, small businesses face full integration since it is irrelevant whether they pay corporate and dividend taxes or pay only personal taxes for the same taxable income.

2.3 The 2006 tax reform

In November 2005, a major dividend tax reform was announced and later implemented in 2006. The main objective of this tax reform was to increase the integration rate for firms that pay the general corporate tax rate. The reform increased the dividend tax credit as well as the gross-up rate faced by shareholders of large corporations (or non CCPCs corporations). Table 2 shows a numerical example of the effects of this reform on taxation.

Table 2: Full integrated tax system after 2006 reform. Active income equal to 1,000.

	General	Small business
Federal corporate tax (-)	221	131
Dividend (net of corporate tax receive)	779	869
Gross up dividend (+)	1000	1000
Personal tax paid (-)	290	290
DTC (+)	185	131
Dividend tax paid	105	159
Dividend received	674	710

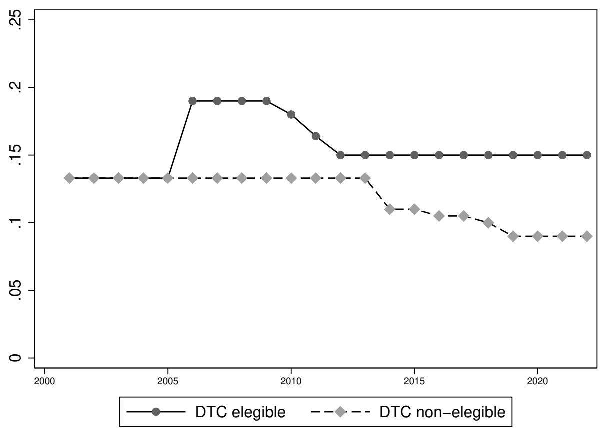
It can be observed by comparing Table 2 with Table 1 that corporate taxable income that pays the general corporate tax rate faced a rise in net dividends received following the reform. This increase was driven by the growth in the gross-up dividend and the increment in the dividend tax credit, which resulted in an additional 2.4% gross corporate active income in the hands of firms owners as net dividends.

2.3.1 Elegible vs Non elegible dividends

The 2006 tax reform incorporated the concept of "eligible" dividends to differentiate between dividends that receive higher dividend tax credits and dividends whose gross corporate taxable income pays the small business tax credit. An "eligible" dividend is paid out of income exceeding the federal threshold and is entitled to a higher DTC, while a "non-eligible" dividend (compensated out of profits below the federal threshold) only receives a DTC such that it matches the small business tax rate.[14]

[14]A tax form was created to differentiate between 'eligible' and 'non-eligible' dividends given by those firms. This form, named Schedule 55, contains the general income rate pool (GRIP). The GRIP registers the amounts of eligible dividends that a CCPC is allowed to issue in a given year. Small firms can only pay eligible dividends up to the amount that appears in the GRIP. The creation of eligible dividends and the GRIP was unannounced, and 'eligible' dividends could be tracked retroactively until 2001 using the GRIP account. If a CCPC gives more eligible dividends than allowed, an additional tax is applied to dividends past the limit.

Figure 4: Dividend tax credit for eligible and non-eligible dividends†



Source: Canadian tax simulator, Milligan (2016).

Figure 4 shows the evolution between dividend tax credits applied to eligible and non eligible dividends. It can noticed from Figure 4 that, after 2005, there is a jump on the dividend tax credit from 13.13% to 18.25%. This higher dividend tax credit translates into a lower dividend tax rate for the shareholder. Theoretically, this increase in the dividend tax credit for taxable income declared after the SBD threshold is an incentive to decrease the number of firms that bunch in the SBD kink. In Section 5, we show the magnitude of the effect of this reform on taxable income declared by firms.

2.4 Data

We use administrative data from the T2 Corporation Income Tax Return Form, which contains tax information for all businesses in Canada. The information is contained at the business number level (line 001 on the T2 tax reform). Additionally, we have access to the firm's province (line 016) and taxable income (line 360).[15] Information on firms' labour inputs is provided by the Longitudinal Employment Analysis Program (LEAP) database, which includes annual employment information for each employer in Canada, starting with the 1983 reference year. The information in LEAP is generated from the annual statements of remuneration paid (T4 slips) that Canadian businesses are required to issue to their employees for tax purposes. The LEAP covers incorporated and unincorporated

[15]Taxable income serves as a proxy for income from active businesses carried on in Canada (line 400).

businesses that issue at least one T4 slip in any given calendar year but excludes self-employed individuals or partnerships where the participants do not draw salaries.

3 The elasticity of taxable income with respect to the dividend tax credit

3.1 Dead weight loss and the elasticity of taxable income

As discussed above, under some assumptions, the elasticity of taxable income is a sufficient parameter to estimate welfare losses associated with changes in the tax rates. In this context, we can extend and complement Devereux et al. (2014) who study the ETI in the context of the corporate tax rate by including the dividend tax credit in an integrated tax system. To do so, we follow Saez et al. (2012) by including the fact that profits at some point are given as dividends; thus, corporate tax revenues will be offset by the dividend tax credit. We also account for the fact that ultimately profits are taxed at the personal tax level. Thus, we can define net of tax profits as:

$$\pi = y - c(y) - T, \tag{1}$$

where y are revenues, $c(y)$ are convex and increasing costs and T are the taxes paid which are defined as:

$$T = \tau^p (1 - \tau^c) (B_c - A_c) + \tau^c (B_c - A_c) - \lambda (B_c - A_c) + E, \tag{2}$$

Where B_c is the relevant tax bracket, A_c is the previous tax base, E are taxes paid by the previous tax brackets, τ^p are personal taxes and λ is the net refundable dividend tax credit. That is, we assume that all profits net of corporate tax are given as dividends.[16] Then, the welfare on this economy is defined by:

$$W = \pi + T \tag{3}$$

[16]This resembles a steady state situation where dividends are eventually paid to firms' owners. We can simply modify this model to incorporate the fact that dividends that are paid in a different period are discounted. In this case, we need to add a discount factor $\xi \equiv \delta_0 + \sum_{i=0}^{\infty} \frac{\delta_{t+s}}{1+r_{t+s}} < 1$, where δ_{t+s} is the proportion of profits from t paid in $t+s$ and r_{t+s} is the respective interest rate. The intuition behind the model does not change after the inclusion of this discount factor.

Now, a change in one minus the dividend tax credit yields:

$$dW = \frac{d\pi}{d(1-\lambda)} \cdot d(1-\lambda) + \frac{dT}{d(1-\lambda)} d(1-\lambda) \tag{4}$$

Using the envelope theorem and accounting for transfers between the firm and the government gives:

$$dW = (\tau^p (1-\tau^c) + \tau^c - \lambda) \frac{\partial B_c}{\partial (1-\lambda)} \cdot d(1-\lambda) = \frac{(\tau^p (1-\tau^c) + \tau^c - \lambda)}{(1-\lambda)} \cdot B_c \cdot \epsilon_{1-\lambda} \cdot d(1-\lambda), \tag{5}$$

with $\epsilon_{1-\lambda} \equiv \frac{\partial B_c}{\partial(1-\lambda)} \cdot \frac{(1-\lambda)}{B_c}$ being the elasticity of taxable income with respect to one minus the dividend tax credit. Intuitively, this parameter should be negative because a higher λ implies a lower tax rate in the hands of firms owners, which incentivizes firms to declare a higher taxable income. In this case, an increase in the absolute value of $\epsilon_{1-\lambda}$ implies a higher welfare increase after an increase in λ. This is driven by the additional taxable income that firms declare because of the increase in λ. In turn, this implies a higher tax base and a higher amount of taxes that the firm and firm owners pay. In addition, following Saez et al. (2012) and Devereux et al. (2014), $\epsilon_{1-\lambda}$, B_c and A_c can be replaced by $\bar{\epsilon}_{1-\lambda}$, $\bar{B}_c$ and $\bar{A}_c$ i.e., the average elasticity of the taxable income, the average tax base and the average previous tax base respectively. Thus, we can write the mechanical effect that ignores any behavioural effect of the dividend tax credit as:

$$dM = \left[\bar{B}_c - \bar{A}_c\right] d(1-\lambda) \tag{6}$$

Thus, the welfare change as a proportion of the mechanical effect is:

$$\frac{dW}{dM} = \frac{\beta \cdot (\tau^p (1-\tau^c) + \tau^c - \lambda)}{\lambda} \cdot \epsilon_\lambda, \tag{7}$$

with $\beta \equiv \dfrac{\bar{B}_c}{\bar{B}_c - \bar{A}_c}$. Assuming that the distribution of $\bar{B}_c$ is Pareto, then β is the shape parameter of a Pareto distribution.[17]

However, the dividend tax credit λ in the Canadian system depends on τ. To see why, first consider the condition on λ for perfect integration:

$$\tau^p = \tau^p (1-\tau^c) + \tau^c - \lambda. \tag{8}$$

[17]This assumption is standard in the literature. See for instance Saez et al. (2012) and Devereux et al. (2014).

Then, λ with perfect integration is equal to:

$$\lambda_p = \tau^c \left(1 - \tau^p\right). \tag{9}$$

Thus, lets assume that

$$\lambda \equiv \alpha \tau^c \left(1 - \tau^p\right),$$

where α is the proportion of the perfect integrated dividend tax credit rate that is given to the shareholder, α is close to 1 when the corporate tax rate is $\tau^c = 0.1313$, and when τ changes to 0.2213, the dividend tax credit is $0.1313 \cdot (1 - \tau^p)$.[18] Knowing this, we can show proposition 1:

Proposition 1. *In an integrated tax system between the dividend and corporate taxes, when the dividend tax credit depends on the corporate tax rate, the elasticity of taxable income with respect to the corporate tax rate is not a sufficient statistic to know welfare changes associated with the corporate tax rate.*

Proof. Changes in welfare derived by changes in the corporate tax rate are given by the following equation:

$$dW = \frac{\partial \pi}{\partial y} \cdot \frac{\partial y}{\partial \left(1 - \tau^c\right)} \cdot d(1 - \tau^c) + \frac{\partial \pi}{\partial \left(1 - \tau^c\right)} \cdot d(1 - \tau^c) + \frac{\partial \pi}{\partial \left(1 - \lambda\right)} \frac{\partial \left(1 - \lambda\right)}{\partial (1 - \tau^c)} \cdot d(1 - \tau^c) + \frac{\partial T}{\partial \left(1 - \tau^c\right)} \cdot d(1 - \tau^c) + \frac{\partial T}{\partial \left(1 - \lambda\right)} \cdot \frac{\partial \left(1 - \lambda\right)}{\partial \left(1 - \tau^c\right)} \cdot d(1 - \tau^c). \tag{10}$$

Now, we can ignore all the transfer effects and apply the envelope theorem, so that we arrive at:

$$dW = \left[\tau^p \cdot (1 - \tau^c) + \tau^c - \lambda\right] \cdot \left[\frac{\partial B_c}{\partial (1 - \tau^c)} + \frac{\partial B_c}{\partial \left(1 - \lambda\right)} \cdot \frac{\partial \left(1 - \lambda\right)}{\partial (1 - \tau^c)}\right] d(1 - \tau^c). \tag{11}$$

Transforming it into elasticities yields:

$$dW = \left[\tau^p \cdot (1 - \tau^c) + \tau^c - \lambda\right] \cdot \left[\frac{\epsilon_{1-\tau^c} \cdot B_c}{(1 - \tau^c)} + \frac{\epsilon_{1-\lambda} \cdot B_c}{(1 - \lambda)} \cdot \frac{\partial \left(1 - \lambda\right)}{\partial (1 - \tau^c)}\right] d(1 - \tau^c) \tag{12}$$

■

[18] $\alpha\left(\tau^c\right)$ can depend on the corporate tax rate. After the tax reform and the threshold, $\tau^c = 0.2313$ and 0.82. This reflect the fact that the integration rate can change when the corporate tax rate changes. However, for the relevant tax bracket B_c it does not change.

This result suggests that when there is perfect integration between the dividend tax credit and the corporate tax, the mechanical effect of rising the corporate tax is zero. That is, a rise in the corporate tax rate does not increase or decrease tax revenues. So, we can obtain the welfare changes relative to the immediate mechanical effect of rising corporate taxes, on the tax base as:

$$dM = \left[- \left(B_c - A_c \right) \right] d(1 - \tau^c). \tag{13}$$

Thus, the welfare change can be written as a proportion of the immediate mechanical effect:

$$\frac{dW}{dM} = -\beta \cdot \left[\tau^p \cdot (1 - \tau^c) + \tau^c - \lambda \right] \cdot \left[\frac{\epsilon_{1-\tau^c}}{(1 - \tau^c)} + \frac{\epsilon_{1-\lambda}}{(1 - \lambda)} \cdot \frac{\partial \left(1 - \lambda \right)}{\partial (1 - \tau^c)} \right]. \tag{14}$$

Where $\epsilon_{1-\tau^c}$ is the elasticity of one minus the taxable income with respect to one minus the corporate tax and $\epsilon_{1-\lambda}$ is the elasticity of one minus the taxable income with respect to the net dividend tax credit. Intuitively, an increase in the corporate tax rate decreases welfare because of the behavioural response. That is, a higher corporate tax incentivizes firm owners to declare a lower amount of taxable income. If this behavioural effect is higher, which is accounted by $\epsilon_{1-\tau^c}$, then the welfare losses are higher after an increase of the corporate tax. Moreover, the behavioural response of the tax base given a change in the dividend tax credit depends on the sign of $\frac{\partial(1-\lambda)}{\partial(1-\tau^c)}$. If this term is positive, then a higher behavioural response induced by an increase of the tax credit is associated with a lower welfare reduction given an increase of the corporate tax rate. The intuition behind this result is that it becomes cheaper for the firm owner to receive dividends; thus, there are incentives to declare more taxable income for the firm. Then $\epsilon_{1-\lambda}$ counters the effect of $\epsilon_{1-\tau^c}$ on the welfare change. The following subsection explains the incentives to bunch that is the basis to estimate both elasticities.

3.2 Conceptual framework and incentives to bunch

To explain bunching incentives, we present a simple stylized model. Assume that firms face heterogeneous convex costs given by $c(y, \theta)$ where θ is a productivity parameter, a higher θ decreases costs. In this case, the after-tax profits for the owners are given by:

$$\pi\left(y,\tau^{p},\tau^{c},\lambda,\theta\right)=y^{\theta}-c(y^{\theta},\theta)-\left(\tau^{p}\left(1-\tau^{c}\right)+\tau^{c}-\lambda\right)\left(B_{c}-A_{c}\right) \tag{15}$$

A firm with a productivity θ that decides to bunch, i.e. decides to declare taxable income just below the SBD threshold K has the following after-tax profits:

$$\pi\left(y,\tau^{p},\tau^{c},\lambda,\theta|\,\text{Bunch}\right)=y^{K,\theta}-c(y^{K,\theta},\theta)-\left(\tau^{p}\left(1-\tau_{1}^{c}\right)+\tau_{1}^{c}-\lambda_{1}\right)\cdot K \tag{16}$$

In this context, the firm with the highest productivity that decides to bunch is indifferent between declaring K as taxable income or declaring $K+\Delta K$ as taxable income. That is, this firm is indifferent between bunching at K or paying a higher marginal corporate tax rate for ΔK. A higher ΔK means that the number of firms that bunch is higher and that the productivity of the highest bunching firm is higher. When this firm decides against bunching, its after-tax profits are:

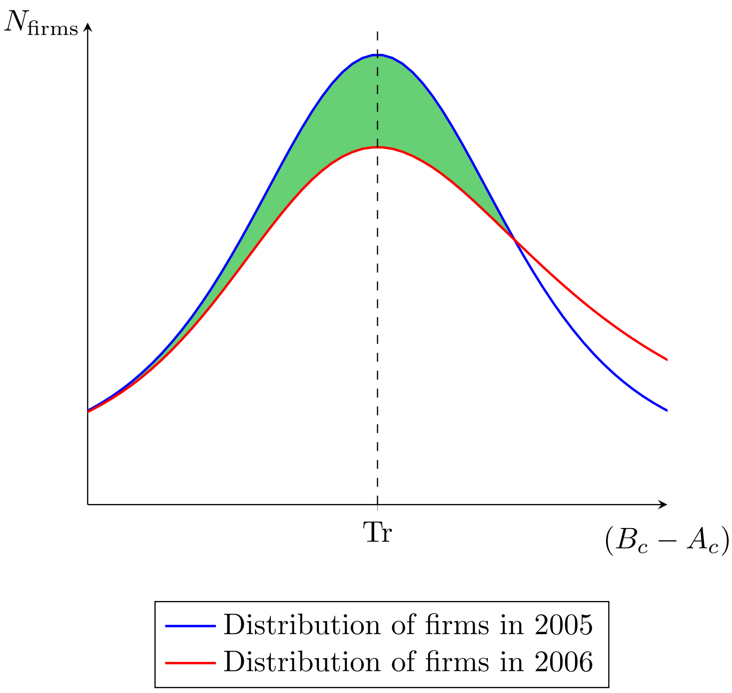
$$\begin{aligned}\pi\left(y,\tau^{p},\tau^{c},\lambda,\theta_{2}|\text{No Bunch}\right)=y^{K+\Delta K,\theta_{2}}-c(y^{K+\Delta K,\theta_{2}},\theta_{2})-\left(\tau^{p}\left(1-\tau_{1}^{c}\right)+\tau_{1}^{c}-\lambda_{1}\right)\cdot K\\ -\left(\tau^{p}\left(1-\tau_{2}^{c}\right)+\tau_{2}^{c}-\lambda_{2}\right)\cdot\Delta K,\end{aligned} \tag{17}$$

where ΔK is the change in taxable income reported by firm θ_2. τ_i^c,λ_i are the corporate tax rate and dividend tax credit rate with bracket $i=1,2$ where 1 is the bracket before the small business deduction (SBD) threshold and 2 is after the SBD threshold. We also know that firm θ_2 is indifferent between declaring K or $K+\Delta K$, which leads to the following condition:

$$\left(y^{K+\Delta K,\theta_{2}}-y^{K,\theta_{2}}\right)-\left(c(y^{K+\Delta K,\theta_{2}},\theta_{2})-\left(c(y^{K,\theta_{2}},\theta_{2})\right)\right)=\left(\tau^{p}\left(1-\tau_{2}^{c}\right)+\tau_{2}^{c}-\lambda_{2}\right)\cdot\Delta K, \tag{18}$$

As can be seen, a higher θ, which is equivalent to a higher ΔK, implies a higher number of firms bunching before the SBD threshold. As suggested by equation (18) two opposite incentives can motivate firms to bunch before the small business deduction (SBD) threshold. First, an increase in the corporate tax rate — which can be seen as an increase in the right-hand side of equation (18) — implies an increase of θ_2 and ΔK. Thus, a higher corporate tax difference motivates more firms to bunch before the threshold. In the Canadian case, the marginal corporate tax rate increases from 13.13% to 23.1% after the threshold. Second, following the 2006 dividend tax credit reform, there was an increase in the dividend tax credit for dividends paid with profits taxed at the higher corporate tax. This increase in the tax credit reduced the incentives for firms to bunch before the SBD threshold. We can see that as a reduction in the right-hand side of equation (18) which implies a decrease of θ_2 and ΔK. Since before 2006 there was no increase in the dividend tax credit, we should expect that the numbers of firms that bunch in 2005 were higher than those bunching in 2006. Figure 5 portrays the anticipated shift in mass between years. Below, we show that we can use the green area, i.e. the difference of bunchers in 2005 and 2006, to estimate the elasticity of taxable income with respect to the dividend tax credit.

Figure 5: Illustrative bunching before the dividend tax reform and after the tax reform.[†]



Note:[†] The y- and x-axis represent the number of firms and the taxable income reported by firms, respectively. The blue line represents the distribution of firms in 2005, before the dividend tax reform, and the red line is the distribution in 2006, after the tax reform. The green area is the difference in the number of bunchers between 2005 and 2006. This image is shown for illustration purpose only and does not depict the actual data.

4 Identification strategy and estimation

In this section, we describe the strategy for identifying the structural parameters as well as the estimation procedure.

4.1 The identification strategy for the elasticity of taxable income with respect to the dividend tax credit

To estimate the 2005 and 2006 bunching parameters, we follow closely the methodology developed in Saez (2010), Chetty et al. (2011), and Devereux et al. (2014). Consider a corporate tax system that introduces a small increase in the marginal corporate tax rate from τ_1 to τ_2 at some income level k. Taxable income below k continues to be taxed at the rate τ_1, and income above k is now taxed at the rate τ_2. In addition, taxable income below k is integrated with the dividend tax credit at a rate λ_1 and income above k is integrated with the dividend tax at a rate λ_2. Under the assumption that there is no effect on firms' actual revenue or costs, the fraction of companies who choose to declare income at the kink point k in response to the small increase in the marginal tax rate can be expressed as:

$$B = B\left(\tau_1^c, \tau_2^c, \lambda_1, \lambda_2\right) = \int_k^{k+\Delta k} g(v)dv, \tag{19}$$

where, $B\left(\tau_1, \tau_2, \lambda_1, \lambda_2\right)$ is the bunching generated by the SBD threshold that includes the corporate tax and dividend tax credit incentives, $g(v)$ is the density of taxable income when τ^c and λ are constant, and $k + \Delta k$ is the highest level of corporate income that a firm with productivity θ_2 could have when it decided to declare taxable income above the kink point. To identify $\epsilon_{1-\lambda}$, the elasticity of taxable income with respect to the dividend tax credit, we need the following set of assumptions:

Assumption 1: We observe bunching before and after the dividend tax credit reform.

Assumption 2: $g(v)$ is uniform around the threshold.

Assumption 3: The corporate tax rate τ^c is the same after the tax reform $\tau^{c\prime}$.

Assumption 4: The SBD threshold K is the same before and after the tax reform K'.

Assumption 5: $g(v)$ is an stationary ditribution, that is $g(v)$ is the same before and after the dividend tax reform.

The following result shows that under Assumptions 1 to 5, ϵ_{λ} is identified.

Proposition 2. *Under Assumptions 1 to 5, $\epsilon_{1-\lambda}$ is identified.*

Proof. First, the numbers of bunchers before the reform is equal to:

$$B\left(\tau_1^c, \tau_2^c, \lambda_1, \lambda_2\right) = \int_k^{k+\Delta k} g(v)dv. \tag{20}$$

And the number of bunchers after the tax reform is equal to:

$$B\left(\tau_1^{c\prime}, \tau_2^{c\prime}, \lambda_1^{\prime}, \lambda_2^{\prime}\right) = \int_{k^{\prime}}^{k^{\prime}+\Delta k^{\prime}} g(v)dv. \tag{21}$$

Using assumptions 1 and 2 yields:

$$B\left(\tau_1^c, \tau_2^c, \lambda_1, \lambda_2\right) = \Delta K g(k). \tag{22}$$

$$B\left(\tau_1^{c\prime}, \tau_2^{c\prime}, \lambda_1^{\prime}, \lambda_2^{\prime}\right) = \Delta K^{\prime} g(K). \tag{23}$$

Now, by assuming that the profit function is isoelastic as in Saez (1999), Chetty et al. (2011) and Devereux et al. (2014) we get:

$$\frac{\Delta K}{K} = \left[\left(\frac{1-\tau_1^c}{1-\tau_2^c}\right)^{\epsilon_{1-\tau^c}} \cdot \left(\frac{1-\lambda_1}{1-\lambda_2}\right)^{\epsilon_{1-\lambda}} - 1\right]. \tag{24}$$

Similarly,

$$\frac{\Delta K^{\prime}}{K^{\prime}} = \left[\left(\frac{1-\tau_1^{c\prime}}{1-\tau_2^{c\prime}}\right)^{\epsilon_{1-\tau^c}} \cdot \left(\frac{1-\lambda_1^{\prime}}{1-\lambda_2^{\prime}}\right)^{\epsilon_{1-\lambda}} - 1\right]. \tag{25}$$

Replacing equations (24) and (25) into (22) and (23) respectively, and using a Taylor expansion around $(\epsilon_{1-\tau^c}, \epsilon_{1-\lambda}) = (0,0)$ gives

$$B\left(\tau_1^c, \tau_2^c, \lambda_1, \lambda_2\right) \approx K \cdot \left[\ln\left(\frac{1-\tau_1^c}{1-\tau_2^c}\right) \cdot \epsilon_{1-\tau^c} + \ln\left(\frac{1-\lambda_1}{1-\lambda_2}\right) \cdot \epsilon_{1-\lambda}\right] \cdot g(K), \tag{26}$$

and

$$B\left(\tau_1^{c\prime}, \tau_2^{c\prime}, \lambda_1^{\prime}, \lambda_2^{\prime}\right) \approx K^{\prime} \cdot \left[\ln\left(\frac{1-\tau_1^{c\prime}}{1-\tau_2^{c\prime}}\right) \cdot \epsilon_{1-\tau^c} + \ln\left(\frac{1-\lambda_1^{\prime}}{1-\lambda_2^{\prime}}\right) \cdot \epsilon_{1-\lambda}\right] \cdot g(K). \tag{27}$$

Subtracting equation (27) to equation (26). and using Assumptions 3 to 5 we have that:

$$\epsilon_{1-\lambda} = \frac{B\left(\tau_1^c, \tau_2^c, \lambda_1, \lambda_2\right) - B\left(\tau_1^{c\prime}, \tau_2^{c\prime}, \lambda_1^\prime, \lambda_2^\prime\right)}{K \cdot g(K) \cdot \ln\left(\frac{1-\lambda_2^\prime}{1-\lambda_2}\right)}. \tag{28}$$

■

Now, define $b\left(\tau_1^c, \tau_2^c, \lambda_1, \lambda_2\right) \equiv \frac{B\left(\tau_1^c, \tau_2^c, \lambda_1, \lambda_2\right)}{g(K)}$, which is the relative bunching at the threshold. Using Assumption 5 yields:

$$\epsilon_{1-\lambda} = \frac{b\left(\tau_1^c, \tau_2^c, \lambda_1, \lambda_2\right) - b\left(\tau_1^{c\prime}, \tau_2^{c\prime}, \lambda_1^\prime, \lambda_2^\prime\right)}{K \cdot \ln\left(\frac{1-\lambda_2^\prime}{1-\lambda_2}\right)}. \tag{29}$$

Hence, $\epsilon_{1-\lambda}$ and $\epsilon_{1-\tau^c}$ can be estimated by finding the relative bunching in the year before the tax reform, the year after the tax reform and with knowledge of the statutory tax parameters. Both elasticities are of interest for policy analysis, given their relationship to the excess burden of taxation and tax avoidance effects. The empirical literature thoroughly discusses the elasticity of taxable income with respect to the corporate income tax. However, to our knowledge, this is the first attempt to estimate the elasticity of taxable income with respect to the dividend tax credit. In the next section, we describe a procedure to estimate such elasticities using the traditional bunching methodology.

4.2 The bunching methodology

Estimating the excess bunching around the SBD threshold requires the imputation of counterfactual distributions. In this regard, we follow Devereux et al. (2014) and Chetty et al. (2011). First, we approximate the distribution with a histogram, dividing the distribution into bins of size κ. Let c_j denote the number of firms in bin j and z_j the mean earnings level relative to the federal threshold (or kink point). The objective is to estimate counterfactual values of c_j. We denote these estimates by $\hat{c}_j$ in the case of bins near the threshold where bunching due to tax planning would be expected to occur. To obtain these estimates, we fit a flexible polynomial of order q to the bin counts in the empirical distribution, excluding bins in the range (z_l, z_u) around the kink point.[19] We choose the

[19]This region is termed the bunching area and does not necessarily have an upper bound below the kink point. It may be optimal for firms engaging in bunching to declare taxable income above the kink point if their marginal cost of incrementally reducing their taxable income is sufficiently high.

bunching area by inspection of the distribution for graphical evidence of bunching. The equation that defines the observed distribution is given by:

$$c_j = \sum_{l=0}^{q} \beta_l z_j^l + \sum_{i=z_l}^{z_u} \gamma_i 1 \left[z_j = j \right] + \sum_{r \in R_k} \rho_{r_k} 1 \left[\frac{z_j}{r} \in N \right] + \epsilon_j, \tag{30}$$

where $\sum_{l=0}^{q} \beta_l z_j^l$ is the polynomial part of the estimation, γ_i is a bin fixed effect for each bin in the bunching region, and $\sum_{r \in R_k} \rho_{r_k} 1 \left[\frac{z_j}{r} \in N \right]$ accounts for rounded-numbers dummies to address the concern that some firms declare income with round numbers and ϵ_j is a stochastic error. The initial estimate of the counterfactual distribution is given by the predicted values from the previous regression conditional on setting all dummies in the excluded range to zero but keeping the contribution of the round-number dummies:

$$\hat{c}_j^0 = \sum_{l=0}^{q} \beta_l z_j^l + \sum_{r \in R_k} \rho_{r_k} 1 \left[\frac{z_j}{r} \in N \right]. \tag{31}$$

The initial estimate of excess bunching, defined as the difference between the observed and counterfactual bin counts within the excluded range, is given by $\hat{B}^0 = \sum_{j=z_l}^{z_u} \left(c_j - c_j^0 \right)$ This value of B^0 is overestimated.[20] To address this issue, we follow Chetty et al. (2011) and Devereux et al. (2014) by shifting the counterfactual distribution to the right of the kink upward. This adjustment is made such that the number of companies in the counterfactual distribution equals the number of companies in the observed distribution. Then, we obtain $\hat{c}_j$ by adjusting the following equation:

$$c_j^{n+1} \left(1 + 1 \left[j > z_u \right] \frac{\hat{B}^n}{\sum_{j=z_u+1}^{\infty} c_j} \right) = \sum_{l=0}^{q} \beta_l z_j^l + \sum_{i=z_l}^{z_u} \gamma_i 1 \left[z_j = j \right] + \sum_{r \in R_k} \rho_{r_k} 1 \left[\frac{z_j}{r} \in N \right] + \epsilon_j. \tag{32}$$

We estimate this equation by iteration until reaching a fixed point, that is when $|B^{n+1} - B^n| \approx 0$. After the fixed point is reached, we can compute $\hat{c}_j$ and B with the estimated parameters, and we can derive the relative excess bunching as:

$$b = \frac{B}{\left(\sum_{j=z_l}^{z_u} \frac{\hat{c}_j}{N} \right)}, \tag{33}$$

[20]This is because the higher tax rate above the kink induces companies above the threshold to decrease their taxable income. Given that the number of firms in each bin tends to fall with taxable income, the observed number of firms in each bin to the right of the kink will tend to be lower than if there were a uniform tax rate (no kink). Hence the estimated counterfactual is likely to underestimate the number of companies that would have been observed in higher-income bins had there been a uniform tax rate.

where N the number of bins in the excluded range. Standard errors are computed using a residual-based bootstrap approach.[21] Additionally, the standard error of the elasticity of taxable income with respect to the under-integration rate is computed using the delta method. Then, we replace our estimates for b in the following equations to obtain the ETI with respect to the corporate tax and the dividend tax credit as:

$$\hat{\epsilon}_{1-\lambda} = \frac{b_{2005} - b_{2006}}{K \cdot \ln\left(\frac{1-\lambda_2^{2006}}{1-\lambda_2^{2005}}\right)}, \tag{34}$$

$$\hat{\epsilon}_{1-\tau^c} = \frac{b_{2005}}{K \cdot \ln\left(\frac{1-\tau_1^{c,2005}}{1-\tau_2^{c,2005}}\right)} - \frac{\ln\left(\frac{1-\lambda_1^{2005}}{1-\lambda_2^{2005}}\right) \cdot \hat{\epsilon}_{1-\lambda}}{\ln\left(\frac{1-\tau_1^{c,2005}}{1-\tau_2^{c,2005}}\right)}, \tag{35}$$

where b_{2005} and b_{2006} represent the relative excess bunching for 2005 and 2006, respectively.

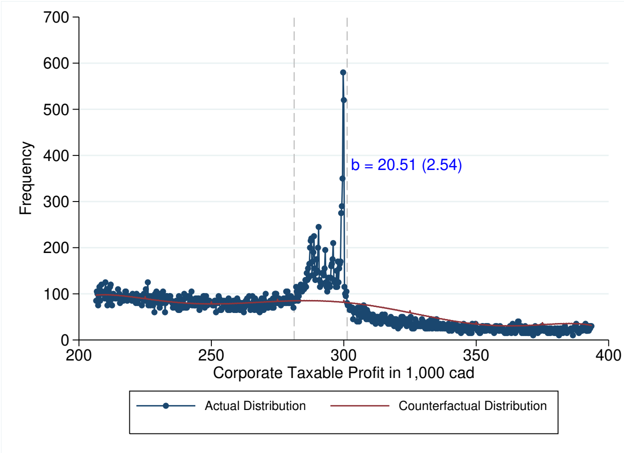
5 Empirical results

5.1 Estimation of excess bunching and the ETI in Canada

We apply the methodology described in the previous section to estimate the excess bunching for Canada during 2005 and 2006. First, we define bins of \$250. Second, by inspection, we set $z_l = \$280,000$ and $z_u = \$304,000$ for 2005. Then, we estimate the counterfactual distribution by following the procedure described above.

[21] From equation (32), we obtain the estimated residual $\hat{e}_j$. We then draw a new set of errors by sampling from the estimated residuals with replacement and create bootstrapped firms counts by adding the new set of errors to the original counts $c_j^h = c_j + \hat{\epsilon}_j^h$. We bootstrap firms' frequencies and follow the same steps above to compute new estimates for the excess mass. We repeat this procedure 500 times and we estimate the standard error as the standard deviation of the 500 estimates.

Figure 6: Distribution of taxable income for 2005.†

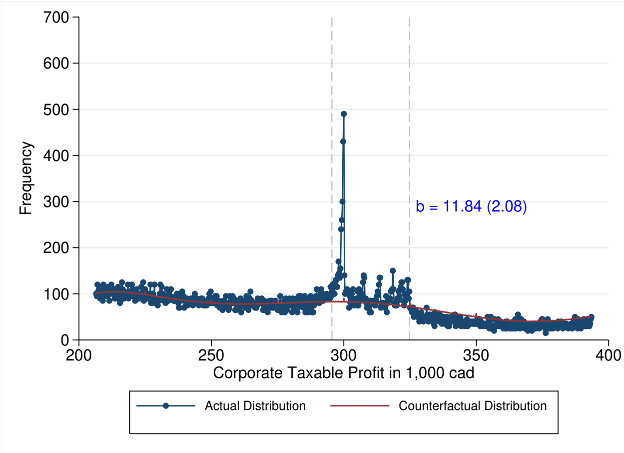


Note†: Bin size is $250. The dashed grey line are the bunching areas and the solid red line is the counterfactual distribution estimated using the methodology from Chetty et al. (2011) and Devereux et al. (2014). b is the relative excess mass. The bootstrapped standard error is reported in parenthesis.

Figure 6 shows the observed and counterfactual densities around $300,000 for 2005. The dashed vertical lines demarcate excluded income ranges, while the solid vertical line represents the $300,000 tax. The blue points plot the observed number of firms in each bin, with the horizontal axis value indicating the upper bound of a given bin. The solid red curve shows the counterfactual distribution based on a fifth-order polynomial using firms with taxable income between $200,000 and $400,000.

There are a few notable observations from the figure. First, there is a large and sharp bunched mass around $300,000. The relative excess mass b_{2005} is estimated to be 20.51 times the density predicted by the counterfactual distribution. This provides strong evidence that firms respond to the tax structure. Second, bunching at $300,000 is asymmetric as the income range that is clearly affected by the bunching around the kink lies between $280,000 and $304,000. Moreover, there is considerably more excess mass to the left of the kink than to its right. Instead, greater mass to the left of the kink appears to reflect some risk aversion in the sense that firms do not want to decrease revenues or increase costs to report revenues too close to the threshold. Figure 7 represent the bunching estimation for 2006.

Figure 7: Distribution of taxable income for 2006.†



Note†: Bin size is $250. The dashed grey line are the bunching areas and the solid red line is the counterfactual distribution estimated using the methodology from Chetty et al. (2011) and Devereux et al. (2014). b is the relative excess mass. The bootstrapped standard error is reported in parenthesis.

Figure 7 shows the observed and counterfactual densities around $300,000 for 2006. Again the excluded income range is demarcated by the dashed vertical lines, again estimated by fitting a fifth-order polynomial for firms with taxable income between $200,000 and $400,000. A few comparisons relative to the figure for 2005 are worth noting. First there is sharp bunching around $300,000, though less than in 2005. Also, although the bunching is still asymmetric, where the bunching windows is between $296,000 and 326,000 now it is more pronounced to the right of the threshold point, which is evidence that there is a shift to the right of the distribution. Also, the relative excess mass b_{2006} is 11.84 times the density predicted by the counterfactual distribution which is close to a half of the relative mass excess found in 2005. This provides evidence that in 2006, firms responded less strongly to the tax structure than in 2005. With this in mind, and using the different tax parameters for 2005 and 2006, we estimate the ETI with respect to the corporate tax and the dividend tax credit using Equations (34) and (35). The results are reported in Table 3.

Table 3 summarizes the information regarding the estimates for 2005, 2006 and the elasticity of taxable income related to the dividend tax credit.[22] It can be observed that the excess bunching attributed to the dividend tax credit is 8.67. With this information,

[22] We use the following parameters to estimate Table 3: $\tau_1^c = 0.1313$, $\tau_2^c = 0.2213$, $\lambda_1 = 0.1313$, $\lambda_2 = 0.185$, $K = \$300,000$.

Table 3: Estimates of the ETI with respect to one minus the corporate tax rate and one minus the dividend tax credit.

Parameter	Point estimate	Increase in 1-MTR	Standard error
$\epsilon_{1-\tau^c}$	0.625	0.109	0.077
$\epsilon_{1-\lambda}$	-0.476	-0.061	0.180

we estimate an ETI with respect to one minus the dividend tax credit equal to -0.476 and an ETI with respect to the corporate tax rate equal to 0.625. These elasticities are 3 to 4 times higher than the values estimated by Devereux et al. (2014) for the UK. This could be explained by the relative excess bunching estimated for the UK, which is in the range of 5.87 to 7.36, whereas for Canada is 20.51 before the reform and 11.84 after the reform. In addition, the tax changes are lower in Canada than in the UK.[23]

5.2 Marginal deadweight cost of an increase of the corporate tax rate in an integrated tax system

We estimate the marginal deadweight cost of the increase in the corporate tax rate before and after the 2006 tax reform. For this, we use equation (14), knowing that $\frac{\partial(1-\lambda)}{\partial(1-\tau^c)}$ is 0 before the 2006 reform. This comes from the fact that before 2006, the integration rate did not change with the SBD threshold. However, after the reform, $\frac{\partial(1-\lambda)}{\partial(1-\tau^c)}$ is equal to 0.82, that is, if the corporate tax increases, the dividend tax credit increases 0.82 of the increase in the corporate tax. Table 4 shows the marginal welfare changes due to the behavioural response of increasing the corporate tax rate, as a proportion of the immediate mechanical effect, assuming that companies with taxable income between \$300,000 and \$400,000 faces a 1 percent increase in their marginal corporate tax rate. Additionally, we estimate β as 1.1.

Table 4: Welfare changes of an increase in the corporate tax rate

	DWL corporate tax before 2006	DWL corporate tax after 2006	Difference
Marginal change	-38%	-15%	23%

Table 4 suggests that, prior to the 2006 reform, an increase in the corporate tax rate reduces welfare by 38% of the immediate mechanical effect. However, following the

[23]A similar explanation can be posited when we compare our results with South Africa. Concretely, Lediga et al. (2019) and Boonzaaier et al. (2019) document elasticities close to those found in Devereux et al. (2014) but with relative excess bunching which is closer to the 2006 Canadian estimation.

dividend tax reform that increased the integration rate, this reduction in welfare is equal to 15 percent. That is, the increase in the integration rate reduces the welfare losses due to increases in the corporate tax rate by 60 percent. This reduction is consistent with the fact that a higher dividend tax credit attenuates the tax planning behaviour before the SBD threshold by incentivizing firms to declare a higher taxable income that will be credited at a higher rate in the hands of the firm owners.

5.3 Robustness checks

5.3.1 Counterfactual polynomial degree

One of the bunching methodology critiques is that the estimates parameters derived from this method are sensitive to the degree of the polynomial chosen to estimate the counterfactual. Because of this, we re-estimate the relative excess mass using polynomials of degree $q \in \{2, 3, 4, 5, 6, 7\}$. Table 5 summarizes the results.

Table 5: Estimates of the relative excess bunching for 2005 and 2006 under alternative counterfactual distributions.†.

	$q = 2$	$q = 3$	$q = 4$	$q = 5$	$q = 6$	$q = 7$
b_{2005}	27.89	26.81	22.88	22.19	20.51	19.44
	(2.64)	(2.46)	(2.64)	(2.38)	(2.38)	(2.58)
b_{2006}	19.01	19.04	14.47	13.7	11.84	11.15
	(2.18)	(2.5)	(1.95)	(2.06)	(2.08)	(2.58)
$b_{2005} - b_{2006}$	8.88	7.77	8.41	8.49	8.67	8.29
	(3.42)	(3.51)	(3.28)	(3.15)	(3.16)	(3.65)

Note†: Estimates of the relative excess bunching for 2005 and 2006 under different functional forms for the counterfactual. q is the polynomial degree used in the counterfactual.

It can be observed that as q increases, b_{2005} and b_{2006} decrease. However, b_{2005} remains greater than b_{2006} and that the difference is statistically different than 0. Additionally, $b_{2005} - b_{2006}$ is stable as q increases.

5.3.2 Identification of the ETI

So far, the estimation of the counterfactual distribution has relied on the assumption of a polynomial functional form, as in Chetty et al. (2011) and Devereux et al. (2014). Nonetheless, Blomquist et al. (2021) criticize the idea of imposing a functional form for the counterfactual density. They argue that the bunching estimator does not identify the taxable earnings elasticity. To understand this, recall equation (36).

$$B = \int_{k}^{k+\Delta z} g(v)dv. \tag{36}$$

Concretely, we only observe the number of people located at the bunching area B, so it is impossible to identify the magnitude of the change in Δz. In turn, this is the parameter used to estimate both the elasticity of taxable income and $g(v)$, which is the counterfactual distribution that induces a technology distribution for firms.[24] Blomquist et al. (2021) argue that assuming a polynomial or a uniform distribution for the counterfactual is stronger than a parametric assumption because it involves imposing an exact form for the firms' technology distribution. To account for this identification issue, we follow Devereux et al. (2014) and Gelber et al. (2020) by estimating a nonparametric counterfactual distribution. The identification assumption for this approach is that the shape of the underlying probability density function is stationary and does not change due to the tax reform.[25] Hence, we use 2009, a post-reform year, where the bunching threshold is found at $\$400,000$. Then, we can directly estimate the counterfactual distribution using the histogram estimator:

$$\hat{p}_H = \frac{c_{j,t_{post}}}{\sum_{i=z_{min}}^{z_{max}} c_{j,t_{post}}}. \tag{37}$$

Then, the counterfactual estimator is:

$$\hat{c}_j = \hat{p}_H \cdot \sum_{i=z_{min}}^{z_{max}} c_{j,t_{pre}}. \tag{38}$$

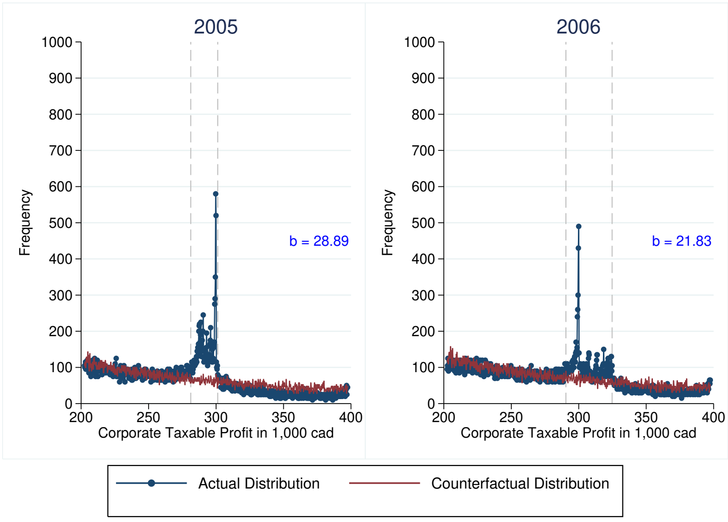
Results are given in Figure 8 in which we can see that the relative excess bunching area is higher when we use a nonparametric bunching. This implies that the ETI with respect to the corporate tax rate is higher. However, the excess bunching difference between 2005 and 2006 is comparable to the baseline estimate (7.06 vs 8.29), which implies that the

[24] Saez (2010) and Chetty et al. (2011) refer to the distribution of preferences over leisure. In the case of firms, there is a technology distribution or a distribution of preferences for tax planning from the owners.

[25] More formally we require that $g(z) = g(z|t)$ where t is time.

elasticity of taxable income with respect to the dividend tax credit is similar.[26]

Figure 8: Distribution of taxable income in 2005 and 2006.†



Note†: Bin size is $250. The dashed grey line are the bunching areas and the solid red line is the counterfactual distribution estimated using a nonparametric estimation based on the 2009 distribution of the taxable income. b is the relative excess mass.

5.4 Heterogeneity analysis

We perform two sets of heterogeneity analysis. One is the difference in bunching in different provinces across Canada and the second is the bunching behaviour of firms that use less than 10 workers.[27]

5.4.1 Provincial analysis

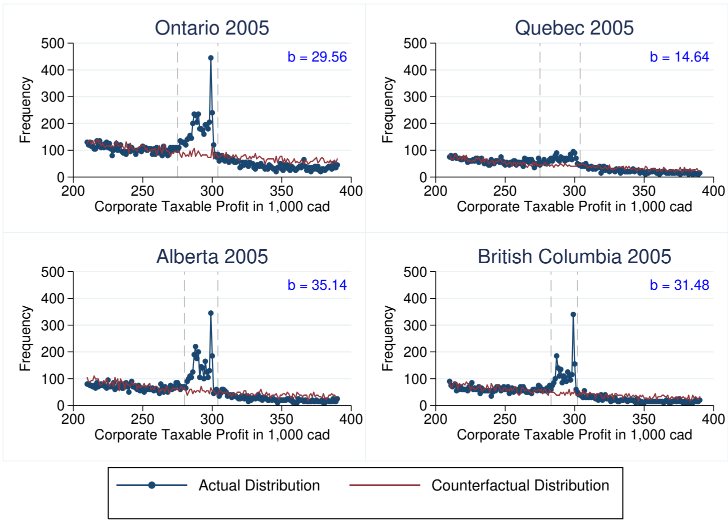
The Canadian tax system is composed of federal and provincial taxes. Differences between provincial codes can incentivise firms to use tax avoidance behaviour in different intensities. Figure 9 shows that firms bunch strongly in Alberta and with less intensity in Quebec. In addition, by comparing Figure 9 with Figure 10, it can be observed see that

[26] Additionally, we estimate the counterfactual distribution by using 2010, 2011 and 2012.

[27] Because of statistics Canada release restrictions, that requires at least ten firms in each bin, we use bins of $1,000 and we estimate this using our nonparametric counterfactual distribution

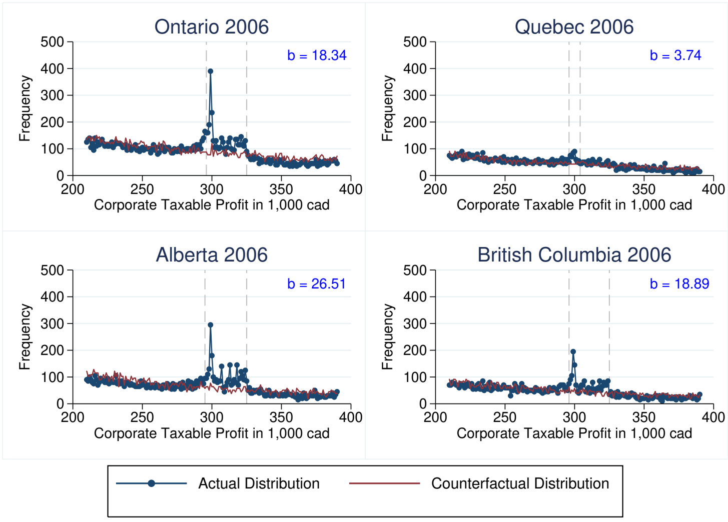
the difference in the response before and after the tax reform is relatively similar as they range from 12.59 in British Columbia to 8.63 in Alberta.

Figure 9: Distribution of taxable income in Ontario, Quebec, Alberta and British Columbia in 2005.†



Note†: Bin size is $1000. The dashed grey line are the bunching areas and the solid red line is the counterfactual distribution estimated using a nonparametric estimation based on the 2009 distribution of the taxable income. b is the relative excess mass.

Figure 10: Distribution of taxable income in Ontario, Quebec, Alberta and British Columbia in 2006.†



Note†: Bin size is $1000. The dashed grey line are the bunching areas and the solid red line is the counterfactual distribution estimated using a nonparametric estimation based on the 2009 distribution of the taxable income. b is the relative excess mass.

5.5 Low workers firms

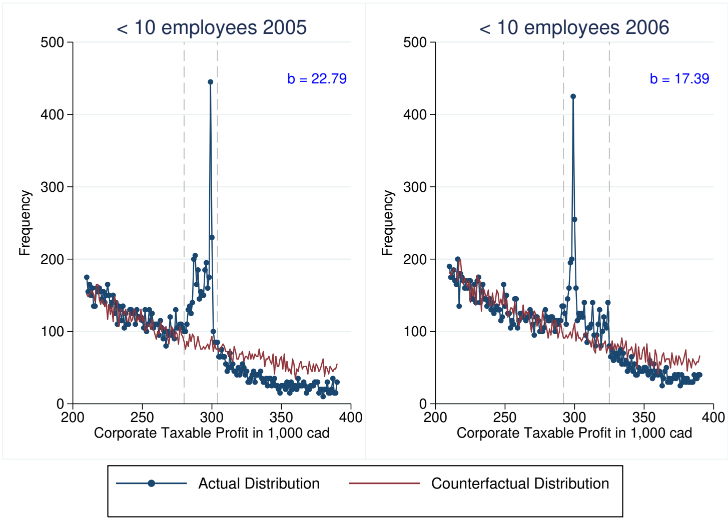
Next, we estimate the excess bunching for firms with less than 10 employees to study heterogeneous bunching in the case of small firms. To measure the number of employees, we use the LEAP dataset.[28] The estimated employment for each enterprise in the LEAP is based on payroll as reported to Canada Revenue Agency. Two measures of employment are available: the Average Labour Unit (ALU) and the Individual Labour Unit (ILU).[29] We use the ALU as a measure of the number of employees. We estimate the bunching behaviour for firms with ALU lower than 10. Figure 11 shows those results. We can see

[28]This universe consists of every business that issued one or more annual statements of remuneration paid for tax purposes (a T4 slip). The self-employed who do not draw a salary are not included in this universe; thus, they are not counted in LEAP. Businesses comprised solely of individuals or partnerships that do not draw a salary are also excluded. The LEAP is a longitudinal file of enterprises, not establishments. The LEAP is constructed using three sources of information: T4 administrative data received from the Canada Revenue Agency, data from Statistics Canada's Business Register (see record no. 1105), and data from Statistics Canada's Survey of Employment, Payrolls and Hours (SEPH, see record no. 2612).

[29]Payroll is converted to ALU using conversion factors derived from the SEPH. The ILU measure is generated using only the payroll information reported by enterprises. The LEAP tracks businesses through time by linking yearly T4 employment data records with enterprise information on Statistics Canada's Business Register.

that the bunching behaviour of those firms is lower than the general universe of firms. In addition, the difference between the excess bunching in 2005 vs 2006 is 5.40, which is lower than the overall population.

Figure 11: Distribution of taxable income in 2005 and 2006 for firms with ALU less than 10 employees.†



Note†: Bin size is $1000. The dashed grey line are the bunching areas and the solid red line is the counterfactual distribution estimated using a nonparametric estimation based on the 2009 distribution of the taxable income. b is the relative excess mass.

6 Policy discussion

In a classical corporate income tax system, shareholders and corporations are taxed as separate units.[30] This approach may cause double taxation of dividends. As noted by Smart (2017), the rationales for integration in Canada were: i) relief from double taxation to reduce the cost of equity finance, ii) investment neutrality between alternative finance instruments, iii) reducing incentives for tax avoidance and iv) promoting investments in Canadian corporations. Thus, integration could increase efficiency by affecting investment, tax avoidance and corporate finance. Indeed, as suggested by our previous estimates, increasing the integration rate improves efficiency by reducing the deadweight loss associated with a one percentage point rise in the tax rate by 60%. That is, the dead-

[30]Or dividends paid to shareholders are not deductible from corporate tax income.

weight loss drops from 38% of the immediate mechanical effect to just 15%. However, there is still tax planning, as Smart (2021) suggested.

In addition, because of tax integration, the increase in the tax credit and the subsequent increase in taxable earnings can be revenue-neutral. Profits that pay a lower corporate tax will receive a lower dividend tax credit but eventually face the same personal tax rate. Thus, if all profits are finally paid as dividends, there is no revenue increase. Nevertheless, there are several ways to avoid personal taxes. In this context, and as the literature has discussed, the corporate tax is a backstop for personal taxes and it is in this backstop that the efficiency gains from the increase of tax integration are generated.

In addition, the cost of increasing the dividend tax credit is not only the mechanical effect but also the cost of raising taxes using another tax base. Moreover, the welfare improvements of the dividend tax credit benefit primarily accrue to the top 1 percent of the income distribution. Thus, the benefit of increasing tax integration should be weighted with the distributional effect of increasing a regressive tax benefit.

7 Conclusion

This work estimates the elasticity of taxable income with respect to the dividend tax credit for Canada. We can see that firms react to the increase in the dividend tax credit by declaring a higher taxable income. We estimate that this elasticity is between 0.4 to 0.5, which is lower than the ETI with respect to the corporate tax rate. The size of both elasticities is an indicator of strong tax avoidance behaviour, which happens near the small business deduction threshold. The increase in the dividend tax credit reduces the deadweight loss of increasing the corporate rate in one percentage point by 60%. As we can see, there are tax avoidance effects derived from the integration between the corporate and dividend tax. Thus, future research should focus on identifying the mechanisms that firms use to reduce their tax payment. For that, it is necessary to link firms' and stakeholders' information so that it becomes possible to know if there is any wage or capital gain increase at the personal level.

References

Bachas, P. and Soto, M. (2021), 'Corporate taxation under weak enforcement', *American Economic Journal: Economic Policy* **13**(4), 36–71.
URL: *https://www.aeaweb.org/articles?id=10.1257/pol.20180564*

Bertanha, M., McCallum, A. H. and Seegert, N. (2021), 'Better bunching, nicer notching', *Working Paper* .
URL: *https://www.federalreserve.gov/econres/feds/files/2021002pap.pdf*

Best, M. C., Brockmeyer, A., Kleven, H. J., Spinnewijn, J. and Waseem, M. (2015), 'Production versus revenue efficiency with limited tax capacity: Theory and evidence from pakistan', *Journal of Political Economy* **123**(6), 1311–1355.
URL: *https://doi.org/10.1086/683849*

Blomquist, S., Newey, W. K., Kumar, A. and Liang, C.-Y. (2021), 'On bunching and identification of the taxable income elasticity', *The Journal of political economy* **129**(8), 2320–2343.

Boonzaaier, W., Harju, J., Matikka, T. and Pirttilä, J. (2019), 'How do small firms respond to tax schedule discontinuities? evidence from south african tax registers', *International tax and public finance* **26**(5), 1104–1136.

Chetty, R. (2009), 'Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods', *Annual Review of Economics* **1**, 451–488.

Chetty, R., Friedman, J. N., Olsen, T. and Pistaferri, L. (2011), ' Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records *', *The Quarterly Journal of Economics* **126**(2), 749–804.

Coles, J. L., Patel, E., Seegert, N. and Smith, M. (2021), 'How do firms respond to corporate taxes?', *Journal of accounting research* .

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

Feldstein, M. (1995), 'The effect of marginal tax rates on taxable income: A panel study of the 1986 tax reform act.', *Journal of Political Economy* **103**, 551–72.

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

Gelber, A. M., Jones, D. and Sacks, D. W. (2020), 'Estimating adjustment frictions using nonlinear budget sets: Method and evidence from the earnings test', *American Economic Journal: Applied Economics* **12**(1), 1–31.
URL: *https://www.aeaweb.org/articles?id=10.1257/app.20170717*

Harberger, A. (1964), *Taxation, Resource Allocation, and Welfare*, Princeton University Press, pp. 25–80.
URL: *http://www.nber.org/chapters/c1873*

Harju, J. and Matikka, T. (2014), Business Owners and Income-Shifting between Tax Bases: Empirical Evidence from a Finnish Tax Reform, CESifo Working Paper Series 5090, CESifo.
URL: *https://ideas.repec.org/p/ces/ceswps/*$_5090.html$

He, D., Peng, L. and Wang, X. (2021), 'Understanding the elasticity of taxable income: A tale of two approaches', *Journal of Public Economics* **197**, 104375.
URL: *https://www.sciencedirect.com/science/article/pii/S0047272721000116*

Kleven, H. J. (2016), 'Bunching', *Annual Review of Economics* **8**(1), 435–464.
URL: *https://doi.org/10.1146/annurev-economics-080315-015234*

Kleven, H. J. and Waseem, M. (2013), 'Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from pakistan', *The Quarterly Journal of Economics* **128**, 669–723.

Kreiner, C. T., Leth-Petersen, S. and Skov, P. E. (2014), 'Year-end tax planning of top management: Evidence from high-frequency payroll data', *American Economic Review* **104**, 154–58.

Kreiner, C. T., Leth-Petersen, S. and Skov, P. E. (2016), 'Tax reforms and intertemporal shifting of wage income: Evidence from danish monthly payroll records', *American Economic Journal: Economic Policy* **8**(3), 233–57.
URL: *https://www.aeaweb.org/articles?id=10.1257/pol.20140233*

le Maire, D. and Schjerning, B. (2013), 'Tax bunching, income shifting and self-employment', *Journal of Public Economics* **107**, 1–18.

Lediga, C., Riedel, N. and Strohmaier, K. (2019), 'The elasticity of corporate taxable incomeevidence from south africa', *Economics letters* **175**, 43–46.

Lobel, F., Scot, T. and Zúniga, P. (2021), 'Corporate taxation and evasion responses: Evidence from a minimum tax in honduras', *Working Paper* .
URL: *https://thiagoscot.github.io/WebsiteDocs/Scot*$_{J}MP_{M}inimumTaxHonduras.pdf$

Piketty, T., Saez, E. and Stantcheva, S. (2014), 'Optimal taxation of top labor incomes: A tale of three elasticities', *American Economic Journal: Economic Policy* **6**, 230–71.

Saez, E. (1999), Do taxpayers bunch at kink points?, Working Paper 7366, National Bureau of Economic Research.
URL: *http://www.nber.org/papers/w7366*

Saez, E. (2010), 'Do taxpayers bunch at kink points?', *American Economic Journal: Economic Policy* **2**, 180–212.

Saez, E., Slemrod, J. and Giertz, S. H. (2012), 'The elasticity of taxable income with respect to marginal tax rates: A critical review', *Journal of Economic Literature* **50**, 3–50.

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

Smart, M. (2017), 'The taxation of dividend income in canada', *Canadian tax journal* **65**(2), 419.

Smart, M. (2021), 'Small business taxation and income inequality: The view from canada', *National tax journal* **74**(2), 493–512.

Waseem, M. (2018), 'Taxes, informality and income shifting: Evidence from a recent pakistani tax reform', *Journal of public economics* **157**, 41–77.