International Tax and Public Finance
https://doi.org/10.1007/s10797-018-9509-0

Using small businesses for individual tax planning: evidence from special tax regimes in Chile

Claudio A. Agostini[1] · **Eduardo Engel**[2] · **Andrea Repetto**[3] · **Damián Vergara**[4]

Abstract
Many countries have special tax regimes (STRs) for small businesses. Even though these regimes may reduce compliance costs, they increase the complexity of the tax system and can be used by high-income individuals to avoid taxes. This paper uses administrative data from Chile to analyze whether the use of STRs is associated with strategic tax planning at the individual level. A descriptive analysis of the data finds three stylized facts that, taken together, are consistent with strategic behavior: STRs are used frequently, they are used mainly by high-income taxpayers, and high-income taxpayers are more likely to hold a portfolio of businesses filing taxes under STRs. We rationalize these facts with a simple model of small business creation and tax planning and test the model's predictions. We find that following a reform that made a particular STR more restrictive, reported individual incomes from businesses filing under that STR decreased between 10 and 15%, while income reported from alternative sources increased. Overall Taxable Income increased between 4 and 7%. This increase is explained by the more restrictive scenario for avoiding taxes through STRs, consistent with individuals using these regimes for tax planning.

Keywords Special tax regimes · Small businesses · Individual tax planning · Tax avoidance · Income sheltering · Behavioral responses

JEL Classification H26 · H25

✉ Claudio A. Agostini
claudio.agostini@uai.cl

1 School of Government, Universidad Adolfo Ibáñez, Santiago, Chile

2 Department of Economics, University of Chile, Santiago, Chile

3 School of Government and Núcleo Milenio Modelos de Crisis (NS 130017), Universidad Adolfo Ibáñez, Santiago, Chile

4 Department of Economics, University of California at Berkeley, Berkeley, USA

Published online: 31 July 2018

1 Introduction

Many countries have special tax regimes (STRs) aimed at reducing the tax compliance costs of small businesses, as well as the tax authority's enforcement costs. Empirical evidence suggests, however, that these special regimes may create incentives for keeping firms small (Engelschalk 2004) and informal (Bird and Wallace 2004; Thuronyi 2004). They can also be used by high-income individuals to avoid taxes (Bird 1974, 1992; Terkper 2003), that is, taxpayers can create small businesses or split large firms into several smaller ones with the sole purpose of sheltering personal income to take advantage of the special tax provisions.

Tax avoidance is an important consideration when designing tax policy (Saez and Stantcheva 2016) because it has implications for horizontal inequity and income inequality, and because it entails efficiency costs (Slemrod and Bakija 2004). If STRs lead to significant levels of tax avoidance, they may overturn the positive effects of these regimes.

This paper uses administrative data from Chile to evaluate whether STRs are used for tax planning decisions at the individual level. We use a novel dataset provided by the Chilean tax authority that includes information from all relevant tax forms filed by taxpayers, allowing us to capture the relation between individual taxpayers and the firms they own under different STRs.

Three stylized facts arise from a descriptive analysis of the data. First, STRs are frequently used: about one in four firms file taxes under one of the main STRs. Second, STRs are mainly used by high-income individuals: over 30% of taxpayers at the top 0.1% of the income distribution own at least one business filing taxes under a STR, compared with only 2.6% of the taxpayers in the bottom 90%. Moreover, almost half of all profits generated by businesses filing taxes under a STR can be attributed to the wealthiest 1%. Third, high-income taxpayers are more likely to hold portfolios of STR firms. The higher the income, the higher the likelihood of holding several businesses subscribed to a STR.[1]

After rationalizing these stylized facts using a simple theoretical model of tax planning and small business creation, we present an econometric analysis that provides evidence of strategic behavior in the use of small businesses and their associated STRs. Our identification strategy exploits a reform that made it more difficult to fulfill the eligibility requirements of one of these regimes: the *Renta Presunta* (Presumptive Tax, PT) regime. This reform was motivated by concerns about its potential use for tax avoidance (Jorratt 2009; Agostini 2013).

Our model predicts that high-income individuals using STRs for tax planning purposes should use them less following a reform that restricted their use. Specifically, incomes reported at the individual level from businesses filing taxes under the PT regime should decrease, while incomes from alternative sources (including business

[1] Even though these stylized facts could reflect the fact that high-income individuals are richer precisely because they have better entrepreneurial abilities and thus create more successful businesses, the analysis below suggests that there may be significant use of STRs for tax avoidance purposes. More important, these stylized facts clearly contradict the goal of STRs as a special provision for low-income owners of small businesses.

income from firms subscribed to other STRs) should increase. Furthermore, overall Taxable Incomes should increase.

We test the model's predictions estimating a difference-in-differences econometric model. We provide evidence that supports the identification assumption (parallel trends) in a highly balanced subsample of high-income taxpayers who earned business income before the reform. Results confirm the model's predictions. The estimation results show that Taxable Incomes of the exposed high-income taxpayers increased between 4 and 7% after the reform relative to the control group. Substitution patterns, as described by the model, lie behind this reaction: while reported incomes from firms filing taxes under the PT regime decreased, reported incomes from other firms (filing taxes under the general regime or under other STRs) and income from independent work increased. These results are consistent with the strategic use of STRs for tax planning at the individual level. They are also robust to several alternative specifications.

This paper contributes to the literature in three dimensions. First, to our knowledge, this is the first paper that studies the process of firm creation and tax regime choice as a strategic tax planning decision at the individual level.

Second, the empirical strategy for estimating behavioral responses is novel: the variation comes from the conditions required to file taxes under STRs and not from the marginal income tax rate, as is usual in this literature.

Finally, recent papers have also found behavioral responses from businesses, especially small ones, to tax policy.[2] For example, by studying the European tonnage taxes, Elschner (2013) show that STRs affect businesses' organizational choice, while Slemrod et al. (2017) show that asking small firms in the USA for additional information about payment card sales increased the likelihood of business income being reported, thereby increasing small businesses' tax compliance. Strategic behavior has also been found regarding eligibility thresholds as they may induce non-compliance behavior (Kanbur and Keen 2014). For instance, Almunia and Lopez-Rodriguez (2018) show that firms in Spain act strategically to avoid stricter tax enforcements triggered by a threshold. In the same line, Onji (2009) shows that large firms in Japan have incentives to masquerade themselves into several small businesses in order to be eligible for VAT tax benefits. As small businesses usually have few owners, firms' behavior may be shaped by their owners' individual strategic decisions. This paper therefore contributes to this literature by proposing a framework that links individuals' and firms' behavioral responses to tax policy.[3]

The remainder of the paper is organized as follows. Section 2 describes the main features of the Chilean income tax system, with a special emphasis on STRs. Section 3 describes the data. Section 4 presents the stylized facts. Section 5 proposes a simple model that accounts for the stylized facts and leads to predictions that are tested in Sect. 6. Finally, Sect. 7 concludes.

[2] This depends on the global complexity of the tax system. For instance, Engelschalk (2004) argues that in transition economies, the existence of STRs has created room for tax evasion and tax avoidance behaviors through the creation of small businesses.

[3] For a survey of individuals' behavioral responses to tax policy, see Saez et al. (2012).

2 The Chilean income tax system

This section describes the main features of the Chilean income tax system as of 2013, the year of the reform that plays a central role in the empirical section of this paper. We emphasize business income and the taxation of shareholder income. In 2013, almost 40% of total tax revenue (6.6% of GDP) was collected through taxes on personal and corporate income.

2.1 Corporate taxes

Under the general corporate tax regime, firms pay a flat 20% tax rate on accrued profits (the *Primera Categoría* regime, henceforth Corporate Tax). In 2013, nearly 72% of active firms filed taxes under this regime (see Sect. 4).

In addition, small businesses are allowed to file taxes under four different STRs. These special provisions intend to reduce the compliance costs faced by small firms and also the enforcement costs of the Chilean tax authority (Jorratt 2013). If tax compliance carries fixed costs for businesses, then small firms bear a disproportional weight when filing taxes. Moreover, burdensome voluntary compliance may push small firms into informality, a risk that may be reduced by tax simplification. In addition, enforcement costs may be reduced by simplified accounting rules. Finally, preferential treatments seek to help small businesses facing financial constraints and other barriers to growth.

The most frequently used STR is the Presumptive Tax regime. About 15% of active firms paid taxes under this regime in 2013. Small firms in three specific economic sectors—agriculture, mining and transport services (both freight and passengers)—are allowed to file taxes under this regime. Taxable Income under the PT regime is the book value of a specific asset or net sales. In the case of agricultural firms, Taxable Income is equal to 10% of the value of the land. In the case of mining firms, Taxable Income is a share of net sales that is increasing in international copper prices. In the case of transport services, taxes are paid over 10% of the value of the vehicles used for transportation.

Before the 2013 reform, a firm in the agricultural or mining sector could benefit from the PT regime if the firm's sales were below a given threshold. This motivated business owners to create several firms with sales just below the threshold so as to file taxes under the PT regime in all of them. Also, in the case of the transport sector, no such threshold existed and all firms were eligible for the PT regime.

The 2013 reform introduced two major changes. First, it associated with every individual a weighted sum of sales of firms in which she participates, the weights equal her participation in the firm.[4] Following the reform, a firm is eligible for the PT regime only if the above-mentioned sums are below the threshold for each one of its owners. If just one owner is ineligible, the firm becomes ineligible to use PT and must file taxes under the general regime. Hence, the incentive to create subsidiaries to avoid taxes was reduced substantially. The reform also introduced a threshold for firms in the transport sector, thereby excluding from the PT regime some large passenger transport companies. Overall, these changes imply that using the PT to avoid taxation became much more difficult following the reform. This will be used to empirically

[4] These sums, and the eligibility criteria that follow, are applied separately for the agricultural, mining and transport sectors.

identify the behavioral response of individual taxpayers to the incentives provided by STRs.

The second most frequently used STR by small businesses is the *14 ter* (henceforth 14T) regime. Under this regime firms paid in 2013 a 20% tax rate on a cash flow basis. Investment and inventories are immediately expensed. This regime simplifies accounting and compliance as it only requires information on sales and expenses. To file taxes as a 14T firm, companies must be VAT taxpayers, have annual sales below 318,000 dollars and register assets below 380,000 dollars. In 2013, 9% of active firms filed taxes under the 14T regime.

Finally, the remaining STRs, *14 bis* and *14 quáter* (henceforth 14B and 14Q, respectively), allow for broader exemptions, although firms have to comply with stricter requirements for eligibility.

Under the 14B regime, small businesses face a zero corporate tax rate and are exempted from keeping detailed accounting and from tracking undistributed profits. Shareholders pay taxes only upon the distribution of profits. To file taxes under 14B, newly created firms have to register capital worth less than 64,000 dollars and annual sales below 445,000 dollars. Active firms can switch to this regime if their annual sales over the past 3 years are on average below 318,000 dollars. A little over 3% of firms filed taxes under this regime in 2013.

Under the 14Q regime, enacted in 2011, firms face a zero corporate tax rate for the first 91,000 dollars of taxable profits. Profits above this amount are taxed under the general corporate tax rules. For eligibility, companies have to sell no more than 1,800,000 dollars annually and register assets under 890,000 dollars. Annual sales for eligibility are computed adding up the sales of all related companies reporting income under 14Q. In 2013, less than 1% of active firms filed taxes under 14Q.

Table 1 summarizes the main features of these five corporate tax regimes.

2.2 Personal income taxes

Labor and pension income are taxed under the progressive *Segunda Categoría* regime (henceforth SC). A Complementary Global Tax, *Global Complementario* (henceforth, GC), applies on all income of individuals, but recognizes as a tax credit the corporate and SC taxes already paid for within the fiscal year. SC and GC income taxes increase progressively through eight identical income brackets with marginal rates between 0 and 40%.[5] It is important to note that the GC tax base only includes distributed profits; that is, retained profits are exempt until distribution. Or, equivalently, corporate taxes cannot be claimed as credit until dividends are paid out. Thus, the corporate tax is simply a withholding tax from personal Taxable Income. Therefore, when profits are distributed among shareholders or owners, the dividends received constitute part of the personal tax base: they are added to all other income, and individuals pay taxes according to the tax bracket in which their income level places them.[6]

[5] Annual income under 12,600 USD is exempt. Nearly 80% of individual taxpayers were exempt in 2013.

[6] Summing up, Taxable Income is equal to total income minus non-distributed profits. That is, Taxable Income equals the sum of dependent and Independent Labor Income dividends, Withdrawals realized capital gains and presumed PT Income. According to Chilean tax law, all dividends and Withdrawals are paid from before-tax business income.

Table 1 Main features of tax regimes for businesses in Chile, 2013

	General corporate tax	14B	14T	14Q	PT
Main benefits	Special treatment for retained profits	Tax base is distributed profit; simplified accounting rules	Tax base is cash flow; simplified accounting rules	Annual profits up to 91,000 dollars are exempt	Income imputed from fiscal value of certain assets; simplified accounting rules
Sales threshold (dollars)	No threshold	318,000 active firms; 445,000 new firms	318,000	1,800,000	From 190,000 to 1,500,000 depending upon the sector
Sales threshold calculated over	–	The firm only	The firm only	All related firms under 14Q	All related firms under PT and within the sector
Assets at registry threshold (dollars)	No threshold	64,000 (new firms)	380,000	890,000	No threshold
Economic sector	All sectors	All sectors	All sectors	All sectors	Agriculture, mining and transport

The exchange rate of December 30, 2013, was used to convert Chilean pesos into dollars

2.3 Discussion

Whenever the tax treatment of income derived from alternative sources differs substantially, taxpayers respond in order to take advantage of the differentials. The STRs analyzed in this paper are likely to have an impact on the way small firms are legally organized and structured, as well as on the way income from different sources is reported at the individual level for tax purposes.

On the one hand, the self-employed may shift labor income into more lightly taxed forms of capital income (e.g., professionals who incorporate can deduct expenses as costs and avoid the higher marginal rates of the SC tax). On the other hand, businesses may disguise one kind of activity as another in order to reduce their tax base and avoid taxes (e.g., create a PT company whose sole investment is a truck and shift profits from other firms to the trucking company). In addition, a firm may split into multiple smaller firms to become eligible for preferential tax treatments (e.g., to take advantage of the tax exemption under 14Q). Finally, personal consumption can be disguised as retained profits to take advantage of the preferential treatment of undistributed profits (e.g., a second or vacation home can be bought as a profit reinvestment under 14B to fully avoid the associated taxes).

Summing up, the STRs described above provide taxpayers with opportunities for tax avoidance and non-compliance. The increased complexity of the income tax system as a whole may also lead to decreased enforcement as it affects the monitoring capacity of the tax authority (Slemrod 1989).

The large gap between the top personal income marginal tax rate (40%) and the corporate flat tax rate (20%) generates incentives for taxpayers for retaining profits, income shifting or consuming profits without distributing. The much lower effective tax rates paid under the different STRs exacerbate those incentives. Therefore, given the progressive schedule of the income tax, high-income taxpayers have the largest incentives to take advantage of the preferential treatments through tax planning strategies.

In the following sections we analyze the tax regime choices made by individual taxpayers in Chile in order to assess whether the observed behavior is consistent with the adoption of tax avoidance strategies, particularly among high-income taxpayers.

3 Data

The data used consist of administrative records covering all Chilean taxpayers, both individuals and businesses, and were provided by the Chilean Internal Revenue Service. We had access to two different datasets. The first consists of data from tax form F22 in the period 2008–2013. This form is used by firms to file corporate income taxes and by individual taxpayers to file personal income taxes (GC). Table 2 shows the number of taxpayers by type filing taxes through form F22.[7]

[7] Dependent workers who do not have income other than wages and salaries do not have to file the F22 tax form, as all their taxes due have been withheld by their employers. Nevertheless, this form is mandatory for individuals perceiving income from firms. On average, 69.8% of all individual taxpayers are exempt from filing this form.

Table 2 Taxpayers filing F22

	Individuals	Firms
2008	1,972,270	912,353
2009	2,010,200	924,870
2010	2,121,840	941,784
2011	2,231,351	967,561
2012	2,732,401	993,418
2013	2,875,947	966,707

The second dataset we use was assembled by the World Bank in collaboration with the Chilean Internal Revenue Service (World Bank 2015). Using administrative data for all taxpayers in 2013, the World Bank merged income from all sources for all individuals (income reported in the F22 form, wages and salaries of dependent workers not filing GC but reported by employers in tax form F1887 and pensions of taxpayers not filing the F22 reported by pension providers in tax form F1812).

Businesses' retained profits were allocated to the corresponding shareholders based on 2013 ownership data. Since many companies in Chile have other companies as shareholders, an iterative process was used by the World Bank to allocate all profits to individuals. The process followed the methods in Agostini (2013) and Fairfield and Jorratt (2015). The stylized facts described in the following section use this dataset to relate individual taxpayers with the activity of the firms they own.

For the estimation of behavioral responses to the PT regime reform, we use the administrative data in all F22 tax forms and focus on high-income individual taxpayers who received income from businesses.[8] One concern is that income generated by PT firms is not observed as it is not used in the calculation of taxes due. To address this concern, we use the tax credits claimed by firms filing taxes under the PT regime as a proxy, a variable that should be highly correlated with PT earnings. Moreover, these tax credits are audited by the Chilean tax authority, so they should represent a reasonably reliable measure of the firms' activity.

Table 3 shows descriptive statistics of the sample used in the benchmark regression analysis. The sample includes all individual taxpayers filing the F22 tax form in 2012 and 2013 who reported business income in 2012 (regardless of the tax regime), were in the highest four tax brackets of the personal income tax schedule,[9] and were 18 years of age or older.

4 STR use: three stylized facts

Three stylized facts regarding the use of STRs are established in this section. First, STRs are frequently used by taxpayers who are business owners. Second, STRs are

[8] Unfortunately, we do not have access to business ownership shares data for years other than 2013. This prevents us from using the broader data used in stylized facts' section in the difference-in-differences estimations below.

[9] Individuals in the highest four tax brackets face marginal income tax rates above the corporate tax rate.

Table 3 Descriptive statistics: average 2012–2013

Statistic	Obs.	Mean	SD
Taxable Income (UF)[a,b]	144,500	2679.43	3519.37
Age (years)	143,510	53.28	13.92
Sex (1 = female)[c]	143,372	0.30	0.46
PT sector[d]	116,968	0.10	0.30
PT Income[e] (UF)	144,500	3.47	38.74
14T income[f] (UF)	144,500	76.36	417.36
Withdrawals[g] (UF)	144,500	889.81	1588.93
Dividends[h] (UF)	144,500	130.34	1084.79
Independent labor income[i] (UF)	144,500	262.96	1212.05
Dependent labor income[j] (UF)	144,500	1035.28	2208.45

[a] All monetary variables are measured in UF., an index that adjusts daily according to last month's CPI. One UF equaled approximately 45 dollars at the end of 2013
[b] Taxable Income is the GC tax base
[c] Sex equals 1 if the taxpayer is female and 0 otherwise
[d] The PT Sector dummy equals 1 if the taxpayer's main economic sector is agriculture, transport or mining, and 0 otherwise
[e] PT Income is the proxy measure for income from firms filing taxes under the PT regime
[f] 14T Income is income from firms filing taxes under the 14T regime
[g] Withdrawals is the income from Withdrawals from firms filing taxes under the 14B, 14Q and the general corporate tax regimes. It is not possible to display statistics for each regime since form F22 pools the three regimes in a unique cell
[h] Dividends is the income from dividends distributed from limited liability companies
[i] Independent Labor Income is income from self-employed labor
[j] Dependent Labor Income consists of pensions, wages and salaries of employees

mainly used by high-income taxpayers. Third, high-income taxpayers hold a portfolio of STR firms. In the next section, we develop a simple model to explain these facts and to obtain predictions that are tested in Sect. 6.

4.1 Stylized fact 1: "STRs are used frequently."

Table 4 shows the fraction of total active businesses filing taxes under each STR in years 2008–2013. About one in four firms file taxes under one of these STRs in any given year. The PT regime is the most frequently used STR: approximately 15% of all active businesses file taxes under this regime. The 14Q regime is the least used regime: less than 1% of all active businesses file taxes under it.

4.2 Stylized fact 2: "STRs are mainly used by high-income taxpayers."

Panel A of Table 5 shows the fraction of taxpayers who own shares of at least one firm filing taxes under any given STR according to the position of the taxpayer in the 2013 income distribution. More than 30% of taxpayers in the top 0.1% own shares of at least one firm filing taxes under a STR, compared with only 2.6% in the bottom

Table 4 Share of total active firms filing taxes under a STR regime. *Source*: Authors' calculations using Chile's Internal Revenue Service data

Year	14B (%)	14T (%)	14Q (%)	PT (%)	Any STR (%)
2008	4.53	5.08	–	15.93	24.42
2009	4.09	5.67	–	15.69	24.48
2010	3.73	6.23	0.10	15.41	24.57
2011	3.38	6.70	0.92	15.09	25.20
2012	3.11	7.29	0.87	14.81	25.20
2013	3.14	9.24	0.96	15.29	27.69

Shares are calculated over total active businesses in each year. The sum of individual STR cells may not exactly match the *Any STR* column, as firms may switch regimes during a year

Table 5 Ownership of STR Firms along the distribution of income. *Source*: Authors' calculations using Chile's Internal Revenue Service data for 2013

Regime	Top 0.1%	0.1–1%	1–5%	5–10%	> 10%
A: fraction of taxpayers owning shares of STR businesses					
14B	5.42%	1.81%	0.73%	0.42%	0.34%
14T	6.76%	3.26%	1.81%	1.18%	0.94%
14Q	2.34%	1.57%	0.83%	0.35%	0.08%
PT	21.24%	10.34%	3.95%	2.50%	1.34%
Any STR	30.71%	15.45%	6.87%	4.21%	2.60%
B: share of profits by income group					
14B	8.74%	27.57%	28.68%	11.76%	23.26%
14T	5.26%	22.93%	27.07%	13.27%	31.47%
14Q	10.04%	36.63%	36.18%	9.18%	7.98%
PT	13.94%	34.41%	24.00%	9.98%	17.67%
Any STR	12.11%	32.27%	25.57%	10.49%	19.55%

Shares in Panel A are calculated over the total number of taxpayers in each income group. Shares in Panel B are calculated over total profits by regime (therefore, each row adds up to 100%)

90%. Differences in ownership along the income distribution are especially strong in the PT regime.

Panel B shows how the share of total profits of businesses filing taxes under the different STRs varies across income groups. Of all profits from STR firms, 44.4% belong to taxpayers in the top 1% (12.1% belong to the top 0.1 and 32.3% to the next 0.9%). This fraction is in sharp contrast with the 19.6% share of profits in the bottom 90%. The PT and the 14Q regimes show the largest concentration of profits at the top, as almost half of total profits under these regimes belong to the richest 1%.

If STRs are used for tax planning purposes, then we should expect a stronger use by high-income individuals: given the progressivity of the income tax schedule, taxpayers at the top face stronger incentives—and also have more resources—to participate in strategic behaviors. Data in Table 5 show that the use of STRs is in fact highly concentrated among the richest individuals. On the one hand, the higher the income,

Table 6 Fraction of taxpayers owning different number of businesses. *Source*: Authors' calculations using Chile's Internal Revenue Service data for 2013

Conditional on owning	0.1%	0.1–1%	1–5%	5–10%	> 10%
A: a single firm					
Any business	5.68%	24.97%	49.72%	66.48%	83.18%
14B	0.77%	12.00%	41.44%	63.57%	87.07%
14T	1.48%	17.67%	47.31%	68.79%	90.10%
14Q	2.49%	15.13%	38.82%	56.32%	72.11%
PT	2.28%	14.34%	41.69%	65.00%	82.07%
B: ten or more firms					
Any business	55.74%	16.75%	4.52%	1.76%	0.70%
14B	77.38%	35.07%	10.07%	3.11%	0.79%
14T	75.71%	24.34%	5.46%	1.85%	0.35%
14Q	63.70%	21.19%	4.95%	1.46%	0.53%
PT	63.24%	21.45%	5.45%	1.51%	0.31%

Shares are calculated over total taxpayers in each income group, conditioning on owning at least one firm subscribed to the corresponding regime

the larger the share of taxpayers owning businesses filing taxes under any given STR. On the other hand, the higher the income, the higher the share of total profits owned by the taxpayers.

4.3 Stylized fact 3: "high-income taxpayers are more likely to hold portfolios of STR firms."

Table 6 shows the fraction of business owner taxpayers at different points of the income distribution who own shares of a single firm (Panel A) or of ten firms or more (Panel B).[10] Panel A shows that 5.7% of individual business owners in the top 0.1% own only one firm, while over 83% of business owners in the bottom 90% do so. In contrast, Panel B shows that 55.7% of the business owners in the top 0.1% own shares of at least 10 firms, but that only 0.7% of business owners at the bottom 90% own that many STR firms.

When we consider individuals owning at least one 14B firm, the fraction of 0.1% top taxpayers who own a single firm drops to less than 1%, while the fraction of individuals owning at least ten firms rises to 77.4%. The reverse is observed when the fraction of taxpayers in the bottom 90% is analyzed. Similar patterns are observed for the other STRs. In summary, when taxpayers are shareholders of a firm filing taxes under any STR, it is likely that the firm belongs to a portfolio of STR firms if the owner is in the upper tail of the income distribution. Conversely, if the taxpayer belongs to the bottom 90%, it is more likely that the firm is the single STR firm owned by the individual.

[10] Table 15 of Appendix B extends this analysis by showing fractions of taxpayers holding shares of 1 to over 15 firms by income group.

Table 7 STR portfolio complexity. *Source*: Authors' calculations using Chile's Internal Revenue Service administrative data for 2013

Regime	0.1%	0.1–1%	1–5%	5–10%	> 10%
A: individuals owning 3 or more firms under any single STR					
14B	6.31%	1.84%	0.40%	0.43%	0.03%
14T	5.80%	2.16%	0.54%	0.24%	0.05%
14Q	4.27%	2.35%	0.58%	0.33%	0.21%
PT	12.39%	5.12%	1.43%	0.54%	0.12%
B: combinations of STR use					
One	84.80%	90.52%	93.56%	94.16%	95.90%
Two	14.14%	9.18%	6.35%	5.78%	4.07%
Three	1.03%	0.29%	0.08%	0.06%	0.03%
All	0.03%	0.01%	0.00%	0.00%	0.00%

Shares in Panel A are calculated over the total number of taxpayers in each income group, conditioning on owning at least one business subscribed to the respective regime. Shares in Panel B are calculated over the total number of taxpayers in each income group, conditional on owning at least one business subscribed to any STR. STR combinations refer to simultaneously owning businesses subscribed to different STRs

Table 7 looks deeper into the ownership of firms subscribed to different STRs. Panel A shows the share of taxpayers who own three or more businesses filing taxes under any given STR, conditional on owning at least one business in that regime. Panel B shows the fraction of taxpayers who simultaneously own businesses subscribed to one or more STR, conditional on owning at least one STR firm.[11]

The figures in Panel A show that the likelihood of owning a large number of STR firms under any given regime is many times higher among the wealthiest groups. More specifically, while 6.3, 5.8, 4.3 and 12.4% of taxpayers at the top 0.1% own three or more businesses subscribed to the 14B, 14T, 14Q and PT regimes, respectively, the corresponding fractions never exceed half a percentage point among the bottom 90%.[12] The same is true for STR combinations: the share of taxpayers simultaneously holding businesses subscribed to more than one special regime increases monotonically with income, from less than 5% at the bottom 90% to around 15% at the top 0.1%.

In summary, this section's descriptive statistics show that STRs are frequently used. More important, they also show that the ownership of these firms is highly concentrated among the wealthiest taxpayers who, at the same time, hold complex portfolios of STR firms. As discussed in the Introduction and Sect. 2, the typical justification for these preferential regimes is that they reduce both compliance costs for small businesses and monitoring costs for the tax authority. Nevertheless, the stylized facts described in this section suggest that more than favoring low-income owners, STR use may be a by-product of the tax planning decisions of high-income individuals.

[11] Table 16 of Appendix B shows detailed data for Panel A, while Table 17 of Appendix B shows detailed data for Panel B.

[12] As Table 16 of'Appendix B shows, the number of businesses in the same regime a taxpayer may own can reach surprisingly high levels. For example, some taxpayers at the top 0.1% own 41 14B businesses, 19 14T businesses or 22 PT businesses.

5 A model of optimal business creation under STRs

In this section we present a simple model where individual taxpayers optimally choose the number of businesses they create under each STR so as maximize their after-tax income. The model accounts for the stylized facts established in Sect. 4 and provides the testable predictions we consider in Sect. 6. A more detailed version of the model is presented in Appendix C.

The individual receives an exogenous income Y and has two STRs at her disposal to reduce her tax burden: regime 1 and regime 2. The individual chooses the number of firms it creates under each regime, S_1 and S_2, to maximize her after-tax income. The sole purpose of creating businesses is to reduce tax liabilities.

We denote by Y_{1s} and Y_{2s} total income sheltered under tax regimes 1 and 2, respectively, and denote their sum by Y_s. Unsheltered income is denoted by Y_u and equal to $Y - Y_{1s} - Y_{2s}$.

Unsheltered income is taxed at a marginal rate of $\tau_m(Y_u)$, with $\tau_m(0) = 0$ and $\tau_m' \geq 0$, and where, for simplicity, we assume τ_m differentiable. The corresponding average rate is denoted as $\tau(Y_u)$.

Income reported by a firm eligible under STR i is taxed at a constant rate t_i. For a firm to be eligible under STR i, reported income can be at most L_i. It follows that, given Y_{is}, the number of firms created under regime i satisfies $S_i = Y_{is}/L_i$, for $i = 1, 2$.

Creating businesses comes at a cost which we assume takes the form:[13]

$$c(S_1, S_2) = c_1(S_1) + c_2(S_2),$$

with $c_i(0) = 0$, $c_i' > 0$ and $c_i'' > 0$ for $i = 1, 2$. Firm creation increases with the number of firms as does the marginal cost of creating a new firm.

We are ready to state the individual's tax planning problem: she solves

$$\begin{aligned} &\max_{Y_u, Y_{1s}, Y_{2s}} [1 - \tau(Y_u)]Y_u + (1 - t_1)Y_{1s} + (1 - t_2)Y_{2s} - c_1(S_1) - c_2(S_2), \\ &\text{s.t.} \quad Y_u + Y_{1s} + Y_{2s} = Y, \\ &S_1 = \frac{Y_{1s}}{L_1}, S_2 = \frac{Y_{2s}}{L_2}, \\ &Y_u \geq 0, Y_{1s} \geq 0, Y_{2s} \geq 0. \end{aligned} \tag{1}$$

The following result rationalizes the stylized facts from Sect. 4.

Result 1 *There exists an income threshold $\overline{Y}$ such that individuals shelter income if and only if $Y > \overline{Y}$. Furthermore, for $Y > \overline{Y}$, the number of businesses created is strictly increasing with income.*

That is, denoting by S_1^* and S_2^* the optimal choices of S_1 and S_2, respectively, there exist income thresholds $\overline{Y}_1$ and $\overline{Y}_2$, both strictly positive, such that $S_i^*(Y) = 0$ if $Y < \overline{Y}_i$ and $dS_i^*(Y)/dY > 0$ for $Y > \overline{Y}_i$. The threshold $\overline{Y}$ mentioned above is equal to $\min(\overline{Y}_1, \overline{Y}_2)$.

[13] This is reasonable if both regimes apply to different economic sectors (as discussed in Sect. 2 for the Chilean case) and ignores economies of scope for businesses that benefit from different regimes.

Proof See Appendix C. □

The above result shows that low-income individuals do not create businesses to benefit from STRs. The reason is that the marginal tax rates they pay on their unsheltered income are low compared to the fixed rate they would have to pay if they reported part of their income as business income. Furthermore, business creation is costly and therefore the income threshold that justifies business creation is higher than it would be in the absence of these costs.

If individuals with higher total income Y do not use STRs, the marginal tax rate they pay on unsheltered income will be higher than the rate they would pay if they reported their marginal dollar as business income. Their optimal strategy, therefore, entails splitting up their total income between unsheltered income and income sheltered in each one of the STRs, in such a way that the after-tax income generated by the marginal dollar, net of business creation costs, is the same for the three decision variables at their disposal. That is:

$$\tau_m(Y_u) = t_1 + \frac{c_1'(Y_{1s}/L_1)}{L_1} = t_2 + \frac{c_2'(Y_{2s}/L_2)}{L_2},$$

with $Y_u + Y_{1s} + Y_{2s} = Y$.

The following result provides predictions for how sheltered income under a particular STR responds when eligibility requirements for the same STR and for a different STR become more demanding. It also provides a prediction for what happens with total sheltered income.

Result 2 *Denote by Y_{1s}^*, Y_{2s}^* the optimal choices for sheltered income under tax regimes 1 and 2, respectively, and denote their sum by Y_s^*. Also denote by Y_u^* the optimal choice of unsheltered income and by T^* total taxes paid by the individual with her optimal portfolio of STR businesses.*

Assume the individual's income, Y, is high enough so that both Y_{1s}^* and Y_{2s}^* are strictly positive. Result 1 provides conditions for this to hold.

Then

$$\frac{\partial Y_{1s}^*}{\partial L_1} > 0, \quad \frac{\partial Y_{2s}^*}{\partial L_1} < 0, \quad \frac{\partial Y_s^*}{\partial L_1} > 0, \quad \frac{\partial Y_u^*}{\partial L_1} < 0, \quad \frac{\partial T^*}{\partial L_1} < 0.$$

That is, a decrease of L_1 that makes regime 1 more restrictive leads to an increase in Y_{2s}^*, Y_u^* and T^* and a decrease in Y_{1s}^* and Y_s^*.

Proof See Appendix C. □

To explain the implications of the above result, assume that regime 1 is the PT regime we described in Sect. 2 and regime 2 represents another STR. The 2013 reform of the PT regime can be captured within our model as a decrease in the threshold L_1 that defines eligibility. Alternatively, as we show in Appendix C, we may replace $c_1(S_1)$ with $a_1 c_1(S_1)$ and study the impact of an increase in a_1.

Result 2 can be interpreted in terms of "income" and "substitution" effects when an STR becomes less attractive for avoidance purposes, while the other STR remains

unchanged. Making the PT regime (that is, regime 1) less attractive leads to a decrease of income sheltered under this regime and an increase of income sheltered under the regime that remained unchanged and therefore has become relatively more attractive. This is the substitution effect. And, since overall STRs have become less effective for tax sheltering, total sheltered income decreases following a decrease in L_1. This is the income effect.

Finally, total taxes paid must increase following a reform that makes eligibility for regime 1 more restrictive. The reason is that the portfolios from which an individual can choose after the reform are a strict subset of the portfolios at her disposal before the reform. Having fewer choices will lead to higher tax payments if the optimal pre-reform portfolio is not an option after the reform, as will be the case for individuals with sufficiently high incomes.

6 Empirical evidence on strategic behavior

On August 2, 2013, the Chilean government sent to Congress a proposal for tax reform that was enacted less than two months later, on September 27. A change in the PT regime was incorporated during the legislative process and ended up being part of the law that was passed. It follows that the change in the PT regime enacted in 2013 can be described as unanticipated.

As discussed in Sect. 2, the reform of the PT regime included two components. First, the introduction of a maximum income level for a business operating in the transport sector to be eligible. This may be interpreted as lowering L_1 in the model described in Sect. 5, where regime 1 corresponds to the PT regime. The reform also consolidated a taxpayer's ownership in all firms within a given sector to determine eligibility. Taken literally, this would amount to imposing $S_1 \leq 1$ in the model. Nevertheless, since there exist ways of partially circumventing this requirement, for example, including spouses, children and relatives among business owners, we may capture this new requirement as an increase in the scale parameter a_1 for the cost function $a_1 c_1(S)$.[14]

In this section, we take advantage of this reform to test the model's predictions. As discussed in the result 2 above a more restrictive scenario for using the PT regime to avoid taxes after the reform should lead to behavioral responses in the form of a portfolio adjustment. Specifically, income reported via firms that benefit from the PT regime may be shifted to other STRs (an increase in Y_{2s} and S_2 in the model), to labor income (an increase in Y_u) and to firms under the general tax regime (given the specifics of the Chilean tax system discussed in Sect. 2 this may be viewed as a combination of an increase in Y_{2s} and Y_u). Any of these responses necessarily increases tax payments if the individual is acting optimally, since all alternative regimes toward which the individual shifts income have either higher tax rates or broader tax bases.

[14] Taxes are filed at the individual level in Chile, that is, there is no joint filing with spouses or other family members.

We test four predictions of the model. First, total taxes paid should increase. Since taxes paid may be contaminated by carryovers to a much larger extent than Taxable Income we actually test whether Taxable Income increases following the reform. Second, income sheltered in businesses that benefit from the PT regime should decrease. Third, income sheltered in businesses that benefit from other preferential regimes, which did not change, should increase. Finally, alternative options available to the taxpayer to generate disposable income, such as Withdrawals from other businesses, dividend payments or Independent Labor Income, should be used more. As discussed above, the four predictions can be interpreted in terms of the model.

6.1 Identification strategy

We use a difference-in-differences (DID) strategy, comparing individual taxpayers who received income from businesses subscribed to the PT regime to taxpayers who obtained income from businesses filing taxes under regimes different from the PT regime, before and after the reform. More specifically, we define treatment and control groups as follows. An individual taxpayer belongs to the treatment group if in the year 2012—the year immediately before the reform—the individual reported income from businesses subscribed to the PT regime. In contrast, the taxpayer belongs to the control group if, even though the individual did report income from firms in 2012, she reported no income from a business under the PT regime that year. Taxpayers with no entrepreneurial activity were excluded from the control group. As the treatment status is not exogenous, sample corrections and complementary robustness checks are carried out below.

The estimated benchmark equation is

$$\log Y_{it} = \alpha + \beta T_i + \delta D_t + \gamma T_i D_t + X'_{it}\theta + \varepsilon_{it}, \tag{2}$$

where $\log Y_{it}$ is the (log of the) outcome variable of interest (Taxable Income or income declared from various sources) reported by individual i in period t; T_i is a dummy variable that indicates whether the individual belongs to the treatment ($T_i = 1$) or control ($T_i = 0$) group; D_t is a dummy variable equal to 1 if $t = 2013$ and 0 otherwise; $T_i D_t$ is the interaction between both variables; X_{it} is a set of control variables that may include age, sex, town and economic sector dummies, and the lagged dependent variable, and ε_{it} is the error term.

The dependent variables used for estimating Eq. (2) are the tax base (Taxable Income), income from businesses filing taxes under the PT regime (PT Income) and Withdrawals from businesses subscribed to the 14B regime, 14Q regime or general regime (Withdrawals),[15] Dividends (i.e., Withdrawals from publicly traded corporations), income from businesses registered as 14T firms (14T Income), income from self-employment (Independent Labor Income) and income from dependent labor (Dependent Labor Income).

[15] The ideal is to use Withdrawals from businesses subscribed to different regimes separately. Yet, as mentioned in Sect. 3, only data on aggregate Withdrawals from these regimes are available.

The coefficient of interest in Eq. (2) is γ, which represents the percentage change of the outcome variable in response to the reform of the PT regime. If businesses subscribed to the PT regime result from income sheltering decisions at the individual level, γ should be positive when the dependent variable is Taxable Income and negative when it is PT Income. The coefficient is also expected to be positive when the dependent variable is income from firms that benefit from STRs other than the PT regime, and when it is income from alternative sources (dividends, Withdrawals and income from self-employment). Finally, we expect γ to be zero (a placebo test) when the dependent variable is income from dependent labor since this usually is not a variable under the taxpayer's control.

Equation (2) is estimated via OLS using panel data to control for changes in the income distribution (Saez 2004). Standard errors are clustered at the individual level. As a robustness check, Eq. (2) is also estimated using a fixed-effects model, a model in differences and a DID model that includes periods previous to 2012 in order to control for previous trends.

The sample is restricted to taxpayers 18 years of age and older in the four highest tax brackets, where individuals face marginal income tax rates above the corporate rate and therefore are more likely to have incentives to use the STRs for avoidance purposes. In terms of the model, these taxpayers are likely to have income above $\overline{Y}$ defined in Result 1.

6.2 Assessing the identification strategy

Since individual taxpayers are not randomly assigned to the treatment and control groups, in this subsection we perform a number of tests assessing the internal validity of our identification strategy. We first assess the balance in covariates and then check for parallel pre-reform trends in the dependent variables.

To assess balance in covariates we follow Imbens and Rubin (2015). We compute four statistics to measure the differences between treatment and control group covariate distributions. Normalized differences, ND_{tc}, are a scale-free statistic that measures differences in distributions' locations. The logarithm of the ratio of standard deviations, Γ_{tc}, measures differences in distributions' dispersion. Finally, the fraction of the treated and control observations whose covariate values are in the tails of the other group's distribution, π_t^{α} and π_c^{α}, are used to assess the supports' overlap.[16] Technical details about the statistics and their empirical implementation ($\widehat{\mathrm{ND}}_{tc}$, $\hat{\Gamma}_{tc}$, $\hat{\pi}_t^{0.05}$ and $\hat{\pi}_c^{0.05}$) are discussed in Appendix D.

As a rule of thumb, values of $\widehat{\mathrm{ND}}_{tc}$ and $\hat{\Gamma}_{tc}$ larger than 0.25, and values of $\hat{\pi}_t^{0.05}$ and $\hat{\pi}_c^{0.05}$ larger than 0.1, may imply sensitivity to specification in linear regression methods.[17] When treatment and control groups are unbalanced, Imbens and Rubin (2015) suggests estimating a propensity score for the treatment status and excluding observations with a probability of being treated that is either too small or too high. Thus, to make our results more reliable, we estimate our specifications excluding

[16] α accounts for the level of confidence, i.e., it defines the *tails* of the distributions. We follow Imbens and Rubin (2015) and use $\alpha = 0.05$.

[17] See Imbens and Wooldridge (2008) and Imbens and Rubin (2015).

Table 8 Assessing balance: four statistics

	Full sample				Trimmed sample			
	$\widehat{\mathrm{ND}}_{tc}$	$\hat{\Gamma}_{tc}$	$\hat{\pi}_C^{0.05}$	$\hat{\pi}_T^{0.05}$	$\widehat{\mathrm{ND}}_{tc}$	$\hat{\Gamma}_{tc}$	$\hat{\pi}_C^{0.05}$	$\hat{\pi}_T^{0.05}$
Taxable Income	0.09	0.14	0.05	0.06	0.07	0.17	0.05	0.05
Lagged Taxable Income	0.11	0.06	0.09	0.06	0.07	0.09	0.07	0.04
Age	**0.27**	− 0.03	0.08	0.05	0.07	0.05	0.04	0.06
Sex (1 = female)	− 0.15	− 0.08	0.01	0.01	0.02	0.02	0.00	0.00
PT sector	**0.69**	**0.69**	**0.24**	**0.20**	**0.64**	**0.52**	0.00	0.00
14T income	0.09	0.12	0.02	0.04	0.06	− 0.08	0.02	0.03
Withdrawals	− 0.14	**0.46**	0.02	0.03	− 0.05	**0.28**	0.02	0.04
Dividends	0.01	− **0.66**	0.03	0.02	0.01	0.22	0.03	0.02
Independent Labor Income	0.03	**0.31**	0.02	0.03	− 0.03	0.01	0.02	0.03
Dependent labor income	0.07	− 0.02	0.02	0.03	− 0.01	− 0.01	0.03	0.02

Values above 0.25 (0.1) in absolute value for the first (last) two statistics are reported in boldface. *Trimmed Sample* stands for the sample that excludes propensity score tails. *Lagged Taxable Income* is the lag of the Taxable Income

observations with propensity score below 0.1 or above 0.9. We also report estimations with the full sample to show that our results are not driven by the sample trimming.

The four statistics are estimated for a large set of variables assumed to be relevant for the analysis. Table 8 shows the statistics for the sample before and after the propensity score correction. All variables are measured in the pre-reform year, i.e., 2012. By construction, the only variables that are unbalanced to an important degree are those related to the PT regime. This is not surprising since PT Income is used to define the treatment status. The overall balance improves considerably after trimming: leaving aside PT-related variables, no covariate shows significant imbalances regarding distribution means, while the overlap statistics show substantial overlap in the central ranges of the distributions. After trimming, only Withdrawals display modest differences in distribution dispersion.

Table 9 presents summary statistics for the full and trimmed samples. Although using propensity scores improves the balance between both groups, differences between the original samples are not large. It follows that the potential loss of external validity associated with using the trimmed sample is unlikely to be of relevance.

We also provide evidence that trends of outcome variables of the treatment and control groups in the absence of treatment are parallel, an assumption needed for the difference-in-differences approach to be valid (Angrist and Pischke 2009). Figure 1 shows the previous trends for Taxable Income before and after the sample trimming. Previous trends for other outcome variables are displayed in Appendix E. Similar to the balance statistics in Table 8, the figures show that the previous trends of most dependent variables are approximately parallel. As an additional check, we estimate Eq. (2) using longer periods in order to control for previous trends and a fixed-effects model. All results are robust to these alternative specifications, consistent with the validity of our identification strategy.

Table 9 Descriptive statistics

	Full sample			Trimmed sample		
	Obs.	Mean	SD	Obs.	Mean	SD
Treatment	144,500	0.15	0.36	78,260	0.20	0.40
Taxable Income (UF)	144,500	2679.43	3519.37	78,260	2684.03	2804.54
Age (years)	143,510	53.28	13.92	78,260	56.80	12.47
Sex (1 = female)	143,372	0.30	0.46	78,260	0.19	0.39
PT Sector	116,968	0.10	0.30	78,260	0.14	0.34
PT Income (UF)	144,500	3.47	38.74	78,260	4.36	27.26
14T income (UF)	144,500	76.36	417.36	78,260	92.12	419.21
Withdrawals (UF)	144,500	889.81	1588.93	78,260	554.92	907.60
Dividends (UF)	144,500	130.34	1084.79	78,260	144.73	688.06
Independent Labor Income (UF)	144,500	262.96	1212.05	78,260	393.79	1430.08
Dependent labor income (UF)	144,500	1035.28	2208.45	78,260	1256.17	2215.58

Trimmed Sample stands for the sample that excludes propensity score tails. UF is an index that adjusts daily according to the CPI. One UF equaled about 45 dollars by the end of 2013

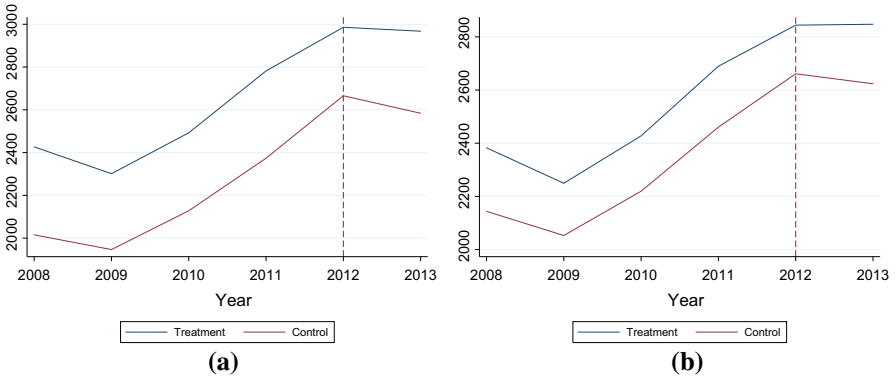


Fig. 1 Previous trends. **a** Taxable Income (all observations). **b** Taxable Income (trimmed sample)

6.3 Results

We begin by estimating the impact of the PT reform on Taxable Income. We estimate (2) with Taxable Income in the role of Y_{it}. Although we prefer the trimmed sample, we present estimations using both samples.

Table 10 reports the estimates of the impact coefficient γ. Columns marked (1) have no controls, columns marked (2) control for sex, age, town and economic sector fixed effects, and columns marked (3) add lagged Taxable Income to the covariates in (2). The results are consistent across specifications. In all cases the estimated coefficient is highly significant, both in economic and in statistical terms, and has the sign predicted by our model. In the preferred specification—trimmed sample with the full set of controls—reported Taxable Income increases by 6.8% as a response to the reform. These results are robust to estimating a fixed-effects model, estimating a model in

Table 10 Main results: Taxable Income

	Full sample			Trimmed sample		
	(1)	(2)	(3)	(1)	(2)	(3)
Taxable Income	0.0697***	0.0699***	0.0956***	0.0462***	0.0465***	0.0680***
	(0.0071)	(0.0079)	(0.0086)	(0.0081)	(0.0081)	(0.0088)
Additional regressors	No	Yes	Yes	No	Yes	Yes
Lagged variables	No	No	Yes	No	No	Yes
Observations	144,500	115,703	114,672	78,260	77,982	77,982

Estimated coefficient γ from (2) with Taxable Income as Y_{it}. Years considered: 2012 (pre-reform) and 2013 (post-reform)
Additional regressors: sex, age, economic sector fixed effects and town fixed effects
Lagged variable: Taxable Income
Standard errors clustered at the individual level reported in parentheses
***10%

Table 11 Main results: income sheltered under different regimes

	Full sample			Trimmed sample		
	(1)	(2)	(3)	(1)	(2)	(3)
PT Income	− 0.0804***	− 0.0851***	− 0.136***	− 0.101***	− 0.0973***	− 0.158***
	(0.00792)	(0.00900)	(0.0128)	(0.00975)	(0.00977)	(0.0139)
14T income	0.142***	0.137***	0.213***	0.158***	0.158***	0.264***
	(0.0151)	(0.0167)	(0.0244)	(0.0184)	(0.0186)	(0.0273)
Additional regressors	No	Yes	Yes	No	Yes	Yes
Lagged variables	No	No	Yes	No	No	Yes
Observations	144,500	115,703	114,672	78,260	77,982	77,982

Estimated coefficient γ from regression (2) using PT Income and 14T income as Y_{it}. Years considered: 2012 (pre-reform) and 2013 (post-reform)
Additional regressors: sex, age, economic sector fixed effects and town fixed effects
Lagged variables: dependent variable and Taxable Income
Standard errors clustered at the individual level reported in parentheses
***10%

differences and considering previous periods for controlling for previous trends. The results of these robustness checks are presented in Appendix F.

Table 11 presents the impact of the reform on income from firms filing taxes under the regime that was reformed (the PT regime) and under a regime that was not reformed (the 14T regime). The model predicts that income sheltered under the PT regime should fall given that this regime became less attractive. The first set of estimates in Table 11 confirms this prediction: income reported from firms filing taxes under the PT regime fell after the reform. For instance, PT Income fell by 15.8% in the preferred estimation. Again, the estimates are consistent across specifications and are robust to alternative estimation approaches (see Appendix F).

The second set of estimates in Table 11 shows the effect of the reform on another STR. This time income reported from firms filing taxes under the 14T regime plays the role of Y_{it} when estimating (2). As predicted by the model, individuals shift

Table 12 Main results: additional options to generate after-tax income

	(1)	(2)	(3)
Withdrawals	0.130***	0.129***	0.270***
	(0.0244)	(0.0248)	(0.0346)
Dividends	0.115***	0.116***	0.170***
	(0.0171)	(0.0174)	(0.0254)
Independent Labor Income	0.0839***	0.0829***	0.105***
	(0.0179)	(0.0181)	(0.0257)
Dependent labor income	− 0.0151	− 0.0171	0.0341*
	(0.0136)	(0.0140)	(0.0186)
Additional regressors	No	Yes	Yes
Lagged variables	No	No	Yes
Observations	78,260	77,982	77,982

Estimated coefficients γ in (2) using Withdrawals, dividends, Independent Labor Income and dependent labor income as Y_{it} and *Trimmed Sample*. Years considered: 2012 (pre-reform) and 2013 (post-reform).
Additional regressors: sex, age, economic sector fixed effects and town fixed effects
Lagged variables: dependent variable and Taxable Income
Standard errors clustered at the individual level reported in parentheses
*1%; ***10%

income reported under the STR that has become less attractive to the STR that now is (relatively) more attractive. Using the trimmed sample and all controls, income from 14T firms increased by 26.4%. Similar results are obtained across specifications and when using alternative estimation procedures (see Appendix F).

Table 12 considers the impact of the PT reform on other margins available to taxpayers to substitute for lower disposable incomes resulting from a more stringent PT regime. The first two sets of estimates consider Withdrawals and dividends obtained from businesses in the role of Y_{it} in (2). The businesses considered include both those under the general tax regime and those filing taxes under STRs other than the 14T regime considered above. As suggested by our model, high-income taxpayers increase the use of alternative sources of income after the PT reform, by 27% in the case of Withdrawals and by 17% in the case of dividends, when we consider the specification with all controls. Again, results are robust to different specifications and econometric modeling assumptions (see Appendix F).

Under the Chilean tax law, business owners can pay themselves honoraria under certain conditions. For tax purposes, these incomes are recorded as Independent Labor Income. The third set of estimates in Table 12 have these earnings as the dependent variable and show that they increased by 10.5% for our preferred specification. By contrast, when we consider Dependent Labor Income, which corresponds mainly to pensions and salaries unrelated to business ownership, the impact of the reform is much smaller and not significant in two out of three specifications. This is consistent with our model, since income from dependent labor cannot be as easily manipulated, in contrast with Independent Labor Income. This explains why we find no significant impact of the reform on dependent labor income reported.

Table 13 Taxable Income in post-reform years

Years	(1)	(2)	(3)
2012–2013	0.0462***	0.0465***	0.0680***
	(0.00807)	(0.00812)	(0.00880)
Observations	78,260	77,982	77,982
2012–2014	0.0344***	0.0369***	0.0472***
	(0.0113)	(0.0113)	(0.0111)
Observations	77,438	77,394	77,153
2012–2015	0.0460***	0.0493***	0.0616***
	(0.0145)	(0.0145)	(0.0135)
Observations	76,236	76,202	75,938
Additional regressors	No	Yes	Yes
Lagged variables	No	No	Yes

Estimated γ in (2) with Taxable Income as Y_{it} and *Trimmed Sample*. Years considered: 2012 (pre-reform) and 2013 (1), 2014 (2) and 2015 (3) as alternative post-reform years
Additional regressors: sex, age, economic sector fixed effects and town fixed effects
Lagged variable: Taxable Income
Standard errors clustered at the individual level reported in parentheses
***10%

Summing up, the empirical results presented in this section are consistent with the four main predictions of a model where individuals strategically set up businesses to maximize their after-tax income. The 2013 reform, which made the PT regime less attractive for tax avoidance, led to an increase in overall Taxable Income of exposed individuals, a decrease in income reported under the reformed regime, an increase in income reported under alternative regimes and no significant change of reported income sources that are largely beyond the taxpayer's control.

6.4 Dynamic effects and heterogeneity

Next we consider two additional results. First, Table 13 presents the medium-term impact of the reform on Taxable Income. To the comparison of post-reform year 2013 with pre-reform year 2012 we considered earlier, we add comparisons of the pre-reform year with the two following post-reform years, 2014 and 2015. The results reported in Table 10 continue holding. In fact, for the latest year available, the magnitude of the effect is very similar to that of the year immediately after the reform: 6.2 versus 6.8%.

Finally, Table 14 estimates heterogeneous impacts of the reform across taxpayers facing different marginal tax rates. Taxpayers in the first three tax brackets face marginal tax rates of 40, 35.5 and 30.4%, respectively, rates that are significantly higher than the corporate tax rate (20%), while taxpayers in the fourth bracket face a much lower marginal tax rate (23%). Estimation results suggest that there is heterogeneity in the behavioral response to the reform: the Taxable Income of taxpayers in the top three brackets increases between 5 and 7%. In contrast, there is no significant effect on the Taxable Income of taxpayers in the fourth bracket. Point estimates show that

Table 14 Heterogeneity by tax bracket

Years	(1)	(2)	(3)
1st Bracket	0.0417**	0.0394*	0.0703***
	(0.0210)	(0.0212)	(0.0225)
Observations	18,602	18,530	18,530
2nd Bracket	0.0530***	0.0521***	0.0531***
	(0.0184)	(0.0186)	(0.0187)
Observations	10,090	10,059	10,059
3rd Bracket	0.0547***	0.0540***	0.0541***
	(0.0151)	(0.0152)	(0.0152)
Observations	16,542	16,487	16,487
4th Bracket	0.0203	0.0203	0.0216
	(0.0164)	(0.0166)	(0.0165)
Observations	18,024	17,952	17,952
Additional regressors	No	Yes	Yes
Lagged variables	No	No	Yes

Estimated γ in (2) with Taxable Income as Y_{it} and *Trimmed Sample*. Years considered: 2012 (pre-reform) and 2013 (post-reform)
Additional regressors: sex, age, economic sector fixed effects and town fixed effects
Lagged variables: lagged Taxable Income
Standard errors clustered at the individual level are reported in parentheses
*1%; **5%; ***10%

the behavioral response is largest among top income taxpayers, consistent with the results of the literature that estimates the elasticity of Taxable Income with respect to marginal tax rates (Saez et al. 2012).

7 Conclusions

This paper analyzes whether the special tax regimes available in Chile for small businesses are used by high-income individuals for strategic tax planning. The paper contributes to the study of firm creation and tax regime choice decisions in the context of tax avoidance.

The empirical evidence based on administrative tax data is consistent with strategic behavior regarding STR use. First, descriptive statistics show that STRs are frequently used, mainly by high-income taxpayers, and that their use appears to be part of a portfolio decision. Second, the econometric analysis supports the hypothesis of tax avoidance: following a reform that made the PT regime stricter, individual Taxable Incomes increased as a result of a portfolio reassignment. In particular, income from firms filing taxes under the PT regime decreased, while other entrepreneurial incomes—both from businesses taxed under the general regime and other STRs—as well as independent labor earnings, increased. Based on the predictions of a simple the-

oretical model that rationalizes the stylized facts, our empirical findings are consistent with the hypothesis of tax planning using STRs at the individual level.

The existence of a significant strategic behavior related to the use of STRs is relevant for the design and evaluation of tax policy both regarding efficiency and equity considerations. Tax avoidance through STRs has efficiency costs as valuable resources and time are spent on artificially creating firms and hiring accountants and tax lawyers (Slemrod and Bakija 2004). Tax avoidance also has implications for horizontal tax equity, as taxpayers with the same income end up facing different tax burdens. It also has implications for after-tax income inequality, as high-income taxpayers have much more to gain from tax avoidance through business creation. The strategic behavior documented in this paper suggests that high-income individuals are more responsive to reforms that restrict the use of STRs for tax avoidance purposes.

Both social preferences and efficiency considerations should be incorporated in the design of optimal tax systems (Saez and Stantcheva 2016). That is, the design, implementation and evaluation of STRs should consider that they can potentially be used for tax planning and tax avoidance purposes. While reducing small firms' compliance costs may be desirable, it is important to assess whether there are alternative policies to the creation of special tax regimes that can address this goal without encouraging tax avoidance.

Acknowledgements We thank Chile's Ministry of Finance and Chile's tax authority, Servicio de Impuestos Internos (SII), for access to the data used in this paper, and Paula Benavides, Conrado Canales, Pamela Castellón, Boris González, Francisco Henríquez, Sandra Luckheide, Eduardo Pantoja and Carlos Recabarren for their help in its analysis. We have benefited from the insights of two anonymous referees, and of Anh Pham, Bibek Adhikari, and seminar audiences at the University of Michigan, the 73rd IIPF Annual Congress, the 110th NTA Annual Conference on Taxation and the 2017 SECHI Annual Meetings. Agostini, Repetto and Vergara acknowledge financial support from CONICYT (DPI20140108). The data provided by the SII's Division of Studies come from individual self-reports and therefore represent only an approximation to actual data. The SII does not take responsibility for the accuracy of the data, nor guarantees its validity or integrity. The responsibility for the analysis and conclusions in this paper lies solely with the authors and does not necessarily represent the views of the Chilean tax administration.

Appendix A: Estimation of financial profits of businesses subscribed to STRs

As businesses subscribed to some STRs are not forced to carry detailed internal accounting, financial profits are not observed. Nevertheless, the Internal Revenue Service of Chile carried out a procedure for estimating financial profits for these businesses in year 2013. In this appendix, we briefly describe the procedure carried out by the Chilean tax authority.

The central assumption is that financial profits are proportional to cash flow. Then, from other forms filed by the businesses, it is possible to compute a cash flow measure for every business i, CF_i, defined by

$$CF_i = S_i - E_i - R_i,$$

where S_i are the sales, E_i are the expenses, and R_i are all wages and salaries paid. This is calculated for all businesses, regardless the tax regime associated, i.e., for the ones taxed by the general scheme and the ones subscribed to STRs.

Consider a set of businesses, A, that do not report profits given they are registered as STR firms. This set is defined by observables (for example, size or economic sector). Then, consider a set of businesses similar in observables, $\hat{A}$, that are taxed under the general regime and, therefore, report information about profits. For those businesses, it is possible to calculate a factor, $F_{\hat{A}}$, from the following relation

$$F_{\hat{A}} = \frac{\sum_{i \in \hat{A}} P_i}{\sum_{i \in \hat{A}} CF_i},$$

where P_i are the profits of firm i in $\hat{A}$. Then, for businesses in A it is possible to estimate profits, P_i, from the following relation

$$P_i = F_{\hat{A}} CF_i, \quad \forall i \in A,$$

i.e., by assuming a proportional relation between profits and cash flow. The groups of businesses taxed by the general regime considered for calculating the factors for the different regimes are

- 14B regime: Businesses with sales under 318,000 USD.
- 14T regime: Businesses with sole proprietorship legal status and sales under 318,000 USD.
- Agricultural PT regime: Businesses of the agricultural sector.
- Mining PT regime: Businesses of the mining sector.
- Freight transportation PT regime: Businesses of the freight transportation sector.
- Passengers transportation PT regime: Businesses of the passengers transportation sector.

Appendix B: Additional tables for stylized fact 3

See Tables 15, 16 and 17.

Table 15 Disaggregation of Table 6. *Source*: Authors' calculations using Chile's Internal Revenue Service data for 2013

Firms	All regimes (%)	14B (%)	14T (%)	14Q (%)	PT (%)
A: 0.1% (97.45%)					
1	5.68	0.77	1.48	2.49	2.28
2	5.64	1.23	2.59	2.85	2.63
3	5.65	1.54	3.08	3.20	4.35
4	5.68	2.31	3.21	4.98	5.18
5	4.46	2.62	2.71	4.63	4.32
6	3.96	2.15	2.47	4.27	3.84
7	3.99	3.69	3.33	4.27	4.43
8	3.60	2.15	2.34	3.91	3.41
9	3.03	3.23	1.60	2.14	3.22
10	2.57	2.92	1.48	3.56	3.10
11	2.72	1.08	2.84	1.78	3.10
12	2.33	1.85	2.22	4.27	2.94
13	2.33	1.23	1.73	4.27	2.67
14	2.25	1.38	1.97	2.85	2.35
15	1.91	1.23	0.99	2.49	2.08
> 15	44.21	70.62	65.97	48.04	50.10
B: 0.1–1% (71.40%)					
1	24.97	12.00	17.67	15.13	14.34
2	17.89	9.04	16.60	14.66	14.85
3	11.63	9.29	11.01	12.60	12.46
4	7.98	8.32	7.43	10.54	9.89
5	5.82	5.87	5.99	8.06	7.99
6	4.34	4.95	5.28	5.65	5.77
7	3.44	5.41	3.69	4.41	4.44
8	2.78	3.62	3.18	2.77	3.55
9	2.30	3.83	2.58	2.59	2.65
10	2.10	2.60	2.24	2.41	2.61
11	1.71	2.55	1.79	2.18	2.18
12	1.49	2.65	1.67	1.82	1.72
13	1.29	2.55	1.70	1.71	1.65
14	1.15	2.45	1.45	1.41	1.62
15	0.95	1.68	1.02	1.18	1.31
>15	10.16	23.18	16.71	12.89	12.97
C: 1–5% (37.38%)					
1	49.72	41.44	47.31	38.82	41.69
2	21.82	19.93	22.09	24.91	23.31
3	9.36	10.52	10.37	12.56	11.99
4	4.95	5.68	4.91	7.10	6.67

Table 15 continued

Firms	All regimes (%)	14B (%)	14T (%)	14Q (%)	PT (%)
5	3.06	4.07	3.28	4.15	3.86
6	2.13	2.61	2.32	2.18	2.39
7	1.57	1.89	1.49	1.75	1.74
8	1.17	1.78	1.23	1.55	1.26
9	0.89	1.12	0.86	1.08	0.87
10	0.83	0.89	0.67	0.95	0.76
11	0.65	0.98	0.59	0.73	0.87
12	0.50	0.98	0.55	0.35	0.62
13	0.41	0.66	0.39	0.58	0.54
14	0.33	0.69	0.41	0.50	0.42
15	0.29	0.49	0.33	0.28	0.45
> 15	2.33	6.28	3.19	2.53	2.56
D: 5–10% (19.77%)					
1	66.48	63.57	68.79	56.32	65.00
2	18.55	19.10	18.52	24.67	20.31
3	5.88	6.70	5.37	8.58	6.48
4	2.69	2.80	2.21	3.96	3.01
5	1.51	1.30	1.05	1.79	1.39
6	1.00	1.14	0.68	1.42	0.82
7	0.77	0.71	0.69	0.80	0.61
8	0.54	0.63	0.31	0.42	0.43
9	0.40	0.47	0.35	0.33	0.30
10	0.42	0.47	0.18	0.24	0.15
11	0.28	0.32	0.20	0.33	0.29
12	0.20	0.16	0.23	0.09	0.17
13	0.17	0.12	0.13	0.14	0.10
14	0.14	0.16	0.13	0.42	0.10
15	0.11	0.12	0.08	0.09	0.10
> 15	0.87	2.24	1.09	0.38	0.75
E: > 10% (9.03%)					
1	83.18	87.07	90.10	72.11	82.07
2	10.68	8.36	7.19	18.66	13.46
3	2.61	2.02	1.33	4.33	2.50
4	1.10	0.65	0.43	1.81	0.79
5	0.60	0.50	0.22	1.25	0.35
6	0.41	0.25	0.17	0.64	0.22
7	0.30	0.13	0.08	0.27	0.12
8	0.18	0.09	0.06	0.23	0.08
9	0.14	0.10	0.03	0.12	0.07
10	0.12	0.04	0.04	0.03	0.04

Table 15 continued

Firms	All regimes (%)	14B (%)	14T (%)	14Q (%)	PT (%)
11	0.09	0.06	0.06	0.08	0.06
12	0.08	0.03	0.03	0.02	0.04
13	0.05	0.02	0.03	0.06	0.02
14	0.03	0.01	0.01	0.02	0.02
15	0.03	0.02	0.01	0.00	0.02
> 15	0.42	0.66	0.21	0.37	0.14

Shares are calculated over total taxpayers in each income group, conditioning of having ownership on at least one business subscribed to the regime specified in the first row. In each panel's title, number in parenthesis shows the share of taxpayers with ownership on at least one business in the income group considered

Table 16 Disaggregation of Panel A of Table 7. *Source*: Authors' calculations using Chile's Internal Revenue Service data for 2013

Regime	1 (%)	2 (%)	3–10 (%)	> 10 (%)	Max
A: 0.1%					
14B	81.54	12.15	4.31	2.00	41
14T	83.97	10.23	5.43	0.37	19
14Q	80.43	15.30	4.27	0.00	6
PT	64.20	23.42	11.68	0.70	22
B: 0.1–1%					
14B	91.07	7.10	1.79	0.05	41
14T	91.38	6.47	2.16	0.00	7
14Q	87.40	10.24	2.35	0.00	8
PT	80.44	14.43	4.95	0.17	41
C: 1–5%					
14B	96.07	3.53	0.40	0.00	7
14T	94.86	4.60	0.54	0.00	6
14Q	93.37	6.05	0.58	0.00	4
PT	90.10	8.47	1.41	0.02	14
D: 5–10%					
14B	97.40	2.17	0.43	0.00	4
14T	96.78	2.98	0.24	0.00	5
14Q	95.47	4.20	0.33	0.00	5
PT	94.32	5.14	0.54	0.00	10
E: > 10%					
14B	99.16	0.81	0.03	0.00	4
14T	98.82	1.13	0.05	0.00	4
14Q	97.42	2.36	0.21	0.00	4
PT	97.74	2.14	0.12	0.00	14

Shares are calculated over total taxpayers in each income group, conditioning on having ownership on at least one business subscribed to the regime specified in the first row. *Max* accounts for the larger value found in the data

Table 17 Disaggregation of Panel B of Table 7. *Source*: Authors' calculations using Chile's Internal Revenue Service data for 2013

Regime	0.1%	0.1–1%	1–5%	5–10%	> 10%
Only 14B	10.64%	8.86%	8.96%	8.50%	11.98%
Only 14T	12.97%	15.04%	21.94%	23.83%	33.19%
Only 14Q	4.94%	8.27%	10.98%	7.80%	3.03%
Only PT	56.26%	58.34%	51.68%	54.03%	47.71%
14B + 14T	1.93%	0.50%	0.24%	0.22%	0.11%
14B + 14Q	0.24%	0.23%	0.16%	0.07%	0.03%
14B + PT	3.77%	1.85%	1.13%	1.20%	1.03%
14T + 14Q	0.14%	0.25%	0.20%	0.11%	0.03%
14T + PT	6.16%	5.08%	3.94%	3.81%	2.76%
14Q + PT	1.90%	1.27%	0.68%	0.37%	0.10%
14B + 14T + 14Q	0.05%	0.01%	0.00%	0.00%	0.00%
14B + 14T + PT	0.73%	0.24%	0.06%	0.04%	0.02%
14B + 14Q + PT	0.24%	0.05%	0.02%	0.02%	0.01%
14T + 14Q + PT	0.00%	0.00%	0.00%	0.00%	0.00%
All	0.03%	0.01%	0.00%	0.00%	0.00%

Shares are calculated over total taxpayers in each income group, conditional on having ownership on at least one business subscribed to a STR. Combinations are interpreted as having ownership shares over businesses subscribed to different STRs

Appendix C: A model of optimal business creation under STRs: detailed version

In this appendix we develop a simple model of individual tax planning in the presence of preferential tax regimes for small businesses. This model leads to the results reported in Sect. 5.

The agent chooses the number of enterprises that benefit from various preferential tax regimes so as to maximize after-tax income. The trade-off she faces when creating a new business is between lowering her tax burden and the cost of setting up and managing the additional enterprise.

We first consider the case with only one tax regime and then extend the model to incorporate two regimes, as in the case considered in Sect. 5. We show that the model's implications are consistent with the stylized facts described in Sect. 4 and derive the implications that are tested in Sect. 6.

Appendix C.1: Model

The model is static. The agent receives exogenous income Y that is strategically broken up into two components, a component that is sheltered from the income tax, Y_s, and an unsheltered component, Y_u, with $Y = Y_s + Y_u$. The unsheltered component pays income tax at marginal rate $\tau_m(Y_u)$; hence, the average income tax rate, $\tau(Y_u)$, satisfies

$$\tau(Y_u) + Y_u \tau'(Y_u) = \tau_m(Y_u). \tag{3}$$

In line with Chile's (and most country's) income tax schedule, we assume $\tau_m(0) = 0$ and $\tau'_m \geq 0$.[18]

The component Y_s is sheltered in small businesses created for the sole purpose of lowering the agent's tax burden. Income reported by each of these businesses up to L is taxed at a constant rate t. Firms have no incentive to report income above L since they loose eligibility for the PT regime should they do so. Consistent with the application in the main text, the special tax regime considered here is referred to as the PT regime.

Creating businesses comes at a cost captured by the function $c(S)$, where S denotes the number of businesses created. These costs can be interpreted as set up costs or the cost of managing the businesses. We assume $c(0) = 0$, $c' > 0$ and $c'' > 0$.[19]

We also assume that these costs cannot be subtracted from the tax base. For simplicity and without loss of generality we ignore integer constraints on S. It then follows that

$$S = \frac{Y_s}{L}, \tag{4}$$

since it is optimal to shelter income in each business to the maximum, L, that benefits from a lower tax rate. It also follows that S is the sum of the agent's participations in all businesses and whether the agent holds entire businesses or only a fraction thereof is irrelevant.

The agent maximizes after-tax income, Z. Given the above assumptions, her problem is

$$\max_{0 \leq Y_s \leq Y} Z = [1 - \tau(Y - Y_s)](Y - Y_s) + (1 - t)Y_s - c(Y_s/L). \tag{5}$$

As shown in the expression above, after-tax income, Z, has two components. The first component is unsheltered income net of income taxes, $(1 - \tau)Y_u$. The second component is sheltered income net of taxes and setup costs, $(1 - t)Y_s - c(S)$.

The above setup captures, albeit in a simplified way, one of the main features of preferential tax regimes for small enterprises described in Sect. 2, namely that their benefit expires beyond a certain size-related threshold. This characteristic provides incentives for high-income individuals to create many such businesses.

Next we solve the agent's problem. Differentiating (5) w.r.t. Y_s and using (3) yields

$$Z'(Y_s) = [\tau_m(Y - Y_s) - t] - \frac{c'(Y_s/L)}{L}. \tag{6}$$

The marginal benefit from creating an additional enterprise is equal to the difference between the gap between both tax rates and the marginal cost of setting up and managing the additional business (where the latter is normalized by the maximum income that benefits from the preferential regime). The first term on the r.h.s. of (6), the difference between both marginal rates, is increasing in sheltered income because the marginal tax rate increases with Taxable Income. The second term, $c'(Y_s/L)/L$, also

[18] For simplicity we assume τ_m differentiable at all points and ignore the possibility of discontinuities.

[19] The cost function c is closely related to a cost of tax avoidance considered in Slemrod and Yitzhaki (2002) and Slemrod (2001) that is increasing and convex in the amount of taxes sheltered.

increases with sheltered income because the marginal cost of creating businesses is increasing. If follows that marginal after-tax income is decreasing in sheltered income:

$$Z''(Y_s) = -\tau_m'(Y - Y_s) - \frac{c''(Y_s/L)}{L^2} < 0. \tag{7}$$

A first implication of (6) and (7) is that the agent will not set up any business if the cost of setting up the first business is larger than the benefit, that is, if $Z'(0) \leq 0$. This leads to

Result C1 *The agent will use the special tax regime only if*

$$[\tau_m(Y) - t]L > c'(0). \tag{8}$$

It follows that there exists a strictly positive income threshold $\overline{Y}$ *characterized as the largest value of* Y *that satisfies*[20]

$$[\tau_m(\overline{Y}) - t]L = c'(0), \tag{9}$$

such that the agent uses the special tax regime only if $Y > \overline{Y}$. *Also, the threshold* $\overline{Y}$ *is increasing in the preferential tax rate* t.

Proof Expression (8) follows from $Z'(Y_s = 0) > 0$. The other statements follow from the assumption that $\tau_m' \geq 0$. □

Agent's problem (5) will have an interior solution if (8) holds and if $Z'(Y_s = Y) < 0$. The latter is equivalent to:

$$[\tau_m(0) - t]L < c'(Y/L)$$

which holds always given the assumption that $\tau_m(0) = 0$. We are ready to characterize the optimal values of sheltered income and the number of businesses:

Result C2 *Consider* $\overline{Y}$ *defined in* (9) *and denote by* Y_s^* *and* S^* *the optimal choices of* Y_s *and* S, *respectively.*

If $Y \leq \overline{Y}$, *we have* $Y_s^* = S^* = 0$. *By contrast, if* $Y > \overline{Y}$, Y_s^* *and* S^*, *are characterized by*[21]

$$[\tau_m(Y - Y_s^*) - t]L = c'(Y_s/L) \tag{10}$$

and

$$[\tau_m(Y - SL) - t]L = c'(S). \tag{11}$$

Proof Follow from (6), (7) and Result C1. □

Result C2 is consistent with stylized facts 1 and 2 in Sect. 4. Special tax regimes will be used by all individuals with income above $\overline{Y}$, with $\overline{Y}$ at least as large as the highest income with an average tax rate of t.

The following result shows that, among those agents that create businesses, the number of businesses held increases with income.

[20] If $\tau_m' > 0$, the value is always unique; otherwise, there may be a continuum of values.

[21] If many values of Y_s^* and S satisfy the conditions that follow, choose the largest one.

Result C3 *Assume the agent's income is larger than* $\overline{Y}$ *defined in* (9)*. Then* Y_s^* *and* S^* *are strictly increasing in* Y*, with*

$$\frac{\partial Y_s^*}{\partial Y} = \frac{\tau_m'(Y_u^*)L^2}{\tau_m'(Y_u^*)L^2 + c''(S^*)} \in [0, 1)\,, \tag{12}$$

$$\frac{\partial S^*}{\partial Y} = \frac{\tau_m'(Y_u^*)L}{\tau_m'(Y_u^*)L^2 + c''(S^*)} \in \left[0, \frac{1}{L}\right), \tag{13}$$

where $Y_u^* = Y - Y_s^*$.

Proof Follows from implicit differentiation of (10) w.r.t. Y. □

The intuition for (12) is the following: as the agent's income increases so does the marginal income tax rate she must pay. For this reason, the agent is prepared to pay higher setup costs when her income is higher. Equation (12) also shows that the marginal propensity to shelter income will lie between zero and one and will be smaller if setup costs grow faster (larger value of c'').

Equation (13) shows that the number of enterprises created will increase with income, as noted in stylized fact 3 in Sect. 4.

The following result complements Result C3 by providing comparative statics w.r.t. variables other than income.

Result C4 *Under the assumptions of Result C3:*

$$\frac{\partial Y_s^*}{\partial t} = -\frac{L^2}{\tau_m'(Y_u^*)L^2 + c''(S^*)} < 0,$$

$$\frac{\partial Y_s^*}{\partial L} = \frac{c'(S^*) + c''(S^*)S^*}{\tau_m'(Y_u^*)L^2 + c''(S^*)} > 0,$$

where $S^* = Y_s^*/L$ *and* $Y_u^* = Y - Y_s^*$*. We also have:*

$$\frac{\partial S^*}{\partial t} = -\frac{L}{\tau_m'(Y_u^*)L^2 + c''(S^*)} < 0,$$

$$\frac{\partial S^*}{\partial L} = \frac{c'(S^*) - \tau_m'(Y_u^*)}{\tau_m'(Y_u^*)L^2 + c''(S^*)}.$$

Finally, to capture changes in the cost of creating businesses, we replace $c(S)$ *by* $ac(S)$*, where* $a > 0$ *is a scale parameter that captures how fast marginal costs increase with the number of firms. We then have:*

$$\frac{\partial Y_s^*}{\partial a} = -\frac{\tau_m'(Y_u^*)L^2 + ac''(S^*)}{c'(S^*)L} < 0, \tag{14}$$

$$\frac{\partial S^*}{\partial a} = -\frac{\tau_m'(Y_u^*)L^2 + ac''(S^*)}{c'(S^*)L^2} < 0. \tag{15}$$

Proof The expressions follow from implicit differentiation of (10) and (11) w.r.t. t, L and a after replacing $c(S)$ by $ac(S)$. Also, to obtain the expressions for partial derivatives w.r.t. L we use (10) to get rid of an expression involving $\tau_m - t$. □

The intuition underlying the first three expressions in Result C4 is straightforward. If the benefits associated with the special tax regime decrease, because t increases or L decreases, sheltered income decreases as well.

In general, the identity $S^* = Y_s^*/L$ implies that the partial derivatives for S^* are obtained dividing partial derivatives for Y_s^* by L. The case of the partial derivative w.r.t. L is different since the denominator in Y_s^*/L also varies with the variable of interest in this case. It is therefore not surprising that the sign of the expression we obtained for $\partial S^*/\partial L$ is indeterminate. If $c'(S^*) > \tau_m'(Y_u^*)$, the optimal number of firms increases with L, while the opposite happens if $c'(S^*) < \tau_m'(Y_u^*)$.

Finally, the intuition for the impact of changes in the cost function is straightforward: a shift upward of this function makes business creation more costly and therefore lowers income sheltered and the optimal number of businesses.

Appendix C.2: The case with two STRs

We extend the above model to the case with two special tax regimes and denote by Y_{is} income sheltered in regime i with $i = 1, 2$ so that now unsheltered income is given by

$$Y_u = Y - Y_{1s} - Y_{2s}.$$

The preferential tax rate of regime i is t_i, valid for reported business income less than L_i, and the number of businesses that benefits from tax regime i is S_i, with

$$S_i = \frac{Y_{is}}{L_i}, \qquad i = 1, 2.$$

As we did in the main text, regime 1 is the PT regime and we will consider the impact of parameter changes in this regime on the agent's choice variables.

We assume two separate cost functions for setting up businesses, one for those of type 1, the other for those of type 2. That is, the cost of setting up and managing S_1 firms of type 1 and S_2 firms of type 2 is

$$c(S_1, S_2) = c_1(S_1) + c_2(S_2) + c_3(S_1 + S_2),$$

with $c_i(0) = 0$, $c_i' > 0$ and $c_i'' > 0$. The first two components are important for the results that follow, that is, both tax regimes involve separate cost components. This will be the case, for example, if both regimes apply to different economic sectors and sector-related sunk investments are needed to be eligible for each regime (see Sect. 2 for examples). The third component captures economies of scope between all businesses where the agent has participation; since its role is not essential to derive the results that follow, we assume it is equal to zero.

Given $Y > 0$, the agent's problem now is:

$$\begin{aligned}\max_{Y_{1s},Y_{2s},Y_u} \quad & Z = [1-\tau(Y_u)]Y_u + (1-t_1)Y_{1s} + (1-t_2)Y_{2s} - c_1(S_1) - c_2(S_2),\\ \text{s.t.} \quad & Y = Y_u + Y_{1s} + Y_{2s},\\ & S_1 = \frac{Y_{1s}}{L_1},\ S_2 = \frac{Y_{2s}}{L_2},\\ & Y_u \geq 0,\ Y_{1s} \geq 0,\ Y_{2s} \geq 0. \end{aligned} \tag{16}$$

Result C5 *Assume the solution to (16) is interior and replace $c_1(S)$ by $a_1 c_1(S)$ in the results with comparative statics involving a_1. Denote the optimal values of Y_{1s}, Y_{2s}, Y_s and Y_u by Y_{1s}^*, Y_{2s}^*, Y_s^* and Y_u^*, respectively. And denote the value of the problem, that is taxes paid under the optimal strategy, by Z^*.*

Then

$$\frac{\partial Y_{2s}^*}{\partial L_1} = -\frac{[c_1''(S_1^*)S_1^* + c_1'(S_1^*)]L_2^2\tau_m'(Y_u^*)}{c_1(S_1^*)c_2(S_2^*) + [c_1''(S_1^*)L_2^2 + c_2''(S_2^*)L_1^2]\tau_m'(Y_u^*)} < 0, \tag{17}$$

$$\frac{\partial Y_s^*}{\partial L_1} = -\frac{c_2''(S_2^*)}{\tau_m'(Y_u^*)L_2^2}\frac{\partial Y_{2s}^*}{\partial L_1} > 0, \tag{18}$$

$$\frac{\partial Y_{2s}^*}{\partial a_1} = \frac{c_1'(S_1^*)L_2^2\tau_m'(Y_u^*)}{c_1''(S_1^*)c_2''(S_2^*) + \left[c_1''(S_1^*)L_2^2 + c_2''(S_2^*)L_1^2\right]\tau_m'(Y_u^*)} > 0, \tag{19}$$

$$\frac{\partial Y_s^*}{\partial a_1} = -\frac{c_2''(S_2^*)}{L_2^2\tau_m'(Y_u^*)}\frac{\partial Y_{2s}^*}{\partial a_1} < 0, \tag{20}$$

where $Y_s^* = Y_{1s}^* + Y_{2s}^*$.

Also, using the identities $Y_{1s}^* = Y_s^* - Y_{2s}^*$ and $Y_u^* = Y - Y_s^*$, the above expressions lead to explicit expressions for Y_{1s}^* and Y_u^* that satisfy $\partial Y_{1s}^*/\partial L_1 > 0$ and $\partial Y_u^*/\partial L_1 < 0$.

Finally,

$$\frac{\mathrm{d}Z^*}{\mathrm{d}L_1} = -c_1'(S_1^*)\frac{S_1^*}{L_1} < 0, \tag{21}$$

$$\frac{\mathrm{d}Z^*}{\mathrm{d}a_1} = c_1(S_1^*) > 0, \tag{22}$$

where $S_1^* = Y_{1s}^*/L_1$.

Proof We derive the partial derivatives w.r.t. a_1; the derivation of partial derivatives w.r.t. L_1 is similar. We also omit the subindex s in Y_{1s} and Y_{2s} in what follows.

From (16) we have that the first-order conditions w.r.t. Y_1 and Y_2 are:

$$\tau_m(Y - Y_1 - Y_2) - t_1 = \frac{a_1}{L_1}c_1'(Y_1/L_1), \tag{23}$$

$$\tau_m(Y - Y_1 - Y_2) - t_2 = \frac{1}{L_2}c_2'(Y_2/L_2). \tag{24}$$

Implicit differentiation of both expressions above w.r.t. a_1, and omitting arguments whenever this does not lead to confusion, leads to:

$$-\tau_m'\left(\frac{\partial Y_1}{\partial a_1}+\frac{\partial Y_2}{\partial a_1}\right)=\frac{c_1'}{L_1}+\frac{c_1''}{L_1^2}\frac{\partial Y_1}{\partial a_1}, \tag{25}$$

$$-\tau_m'\left(\frac{\partial Y_1}{\partial a_1}+\frac{\partial Y_2}{\partial a_1}\right)=\frac{c_2''}{L_2^2}\frac{\partial Y_2}{\partial a_1}. \tag{26}$$

Subtracting (26) from (25) yields:

$$\frac{\partial Y_1}{\partial a_1}=\frac{c_2''L_1^2}{c_1''L_2^2}\frac{\partial Y_2}{\partial a_1}-\frac{c_1'L_1}{c_1''}.$$

Substituting this expression for $\partial Y_1/\partial a_1$ in (26) and solving for $\partial Y_2/\partial a_1$ leads to (19). And substituting the expression for $\partial Y_2/\partial a_1$ from (19) in (24) yields (20).

Finally, (21) and (22) follow from the envelope theorem. To see this, we write the taxpayer's problem substituting the constraints into the objective function:

$$\begin{aligned}&\max_{Y_{1s},Y_{2s}}[1-\tau(Y-Y_{1s}-Y_{2s})](Y-Y_{1s}-Y_{2s})+(1-t_1)Y_{1s}\\&\quad+(1-t_2)Y_{2s}-a_1c_1(Y_{1s}/L_1)-c_2(Y_{2s}/L_2).\end{aligned}$$

This concludes the proof. □

As stressed in Sect. 5, Result C5 establishes what may be viewed as "income" and "substitution" effects when a preferential tax regime—the PT regime in the case of this paper—becomes less attractive, either because L_1 decreases or because a_1 increases. In both cases total sheltered income decreases, which follows from (18) and (20). This is the income effect and is closely related to the fact that the agent is poorer. At the same time, the individual switches sheltered income from the regime that became less attractive to the one unaffected by the reform, as shown by (17) and (19). This is the substitution effect.

Appendix C.3: Proof of results in Sect. 5

Result 1 in Sect. 5 is a straightforward extension of Results C1, C2 and C3 to the case of two STRs. Result 2 follows from Result C5.

Appendix D: Statistics for assessing balance

In this section, details about the statistics proposed by Imbens and Rubin (2015) for assessing balance in covariates are discussed. The first one, *normalized differences*, is a scale-free way for measuring the difference in locations of the distributions. It is defined by

$$\mathrm{ND}_{tc}=\frac{\mu_t-\mu_c}{\sqrt{\left(\sigma_t^2+\sigma_c^2\right)/2}},$$

where t and c denote treatment and control groups, respectively, and (μ_i,σ_i^2) are the population mean and variance of group i, for $i=t,c$, of a given variable X. This measure can be empirically implemented by

$$\widehat{\mathrm{ND}}_{tc} = \frac{\bar{X}_t - \bar{X}_c}{\sqrt{\left(s_t^2 + s_c^2\right)/2}},$$

where $\bar{X}_i = \frac{1}{N_i}\sum_{j\in i} X_j$ and $s_i^2 = \frac{1}{N_i - 1}\sum_{j\in i}\left(X_j - \bar{X}_i\right)^2$, with N_i denoting the number of observations belonging to group i, for $i = t, c$. Imbens and Rubin (2015) suggest that $\widehat{\mathrm{ND}_{\mathrm{tc}}}$ is better than the t-statistic for assessing differences in distributions. The central idea behind assessing balance is not to determine whether there is enough information about differences in covariate means, but to analyze whether or not differences are large enough to invalidate a posterior econometric application. The scale-free nature of the statistic is beneficial for that purposes.

For assessing differences in distributions' dispersion, the authors propose the use of the logarithm of the ratio of standard deviations,

$$\Gamma_{tc} = \ln\left(\frac{\sigma_t}{\sigma_c}\right) = \ln(\sigma_t) - \ln(\sigma_c),$$

which can be empirically implemented by

$$\hat{\Gamma}_{tc} = \ln(s_t) - \ln(s_c).$$

Finally, the analysis can be complemented by calculating the fraction of treated and control observations whose covariate values are in the tails of the other group's distribution. The idea is to determine whether the comparison between units of the different groups will rely too much on extrapolation. Fixing a confidence value α, the probability mass that is outside the tails of the other group's distribution is

$$\pi_i^{\alpha} = \left(1 - F_i\left(F_j^{-1}(1-\alpha/2)\right)\right) + F_i\left(F_j^{-1}(\alpha/2)\right),$$

where F_i is the cumulative distribution function for $i = t, c$ and j is the other group. With F unknown, this statistic can be empirically implemented using the empirical distribution functions

$$\hat{F}_i(x) = \frac{1}{N_i}\sum_{j\in i} 1_{X_j \le x},$$

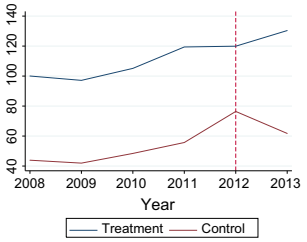
where $1_{X_j \le x}$ is an indicator variable that takes value 1 if $X_j \le x$, and

$$\hat{F}_i^{-1}(q) = \min_{-\infty < x < \infty}\left\{x : \hat{F}_i(x) \ge q\right\},$$

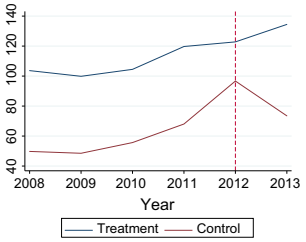
for $i = t, c$. Then, fixing $\alpha = 0.05$, statistics can be empirically implemented by

$$\hat{\pi}_i^{0.05} = \left(1 - \hat{F}_i\left(\hat{F}_j^{-1}(0.975)\right)\right) + \hat{F}_i\left(\hat{F}_j^{-1}(0.025)\right).$$

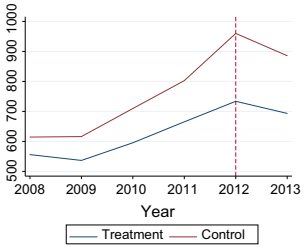
Appendix E: Previous trends: other variables



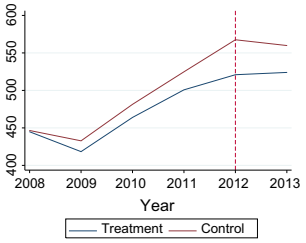
(a) 14T Income (all observations)



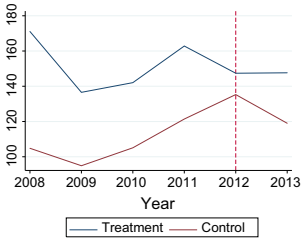
(b) 14T (trimmed sample)



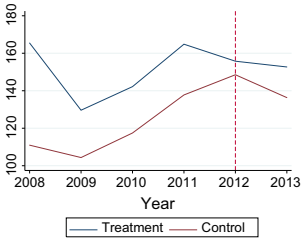
(c) Withdrawals (all observations)



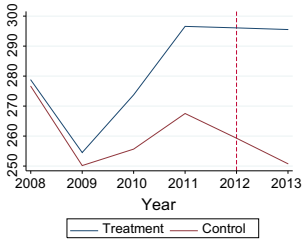
(d) Withdrawals (trimmed sample)



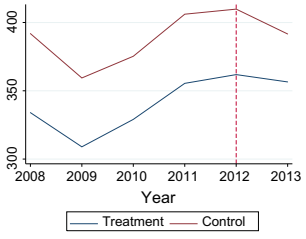
(e) Dividends (all observations)



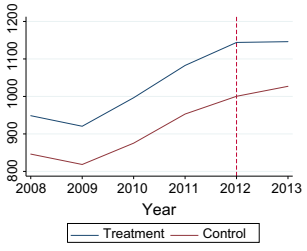
(f) Dividends (trimmed sample)



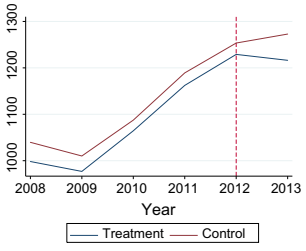
(g) Indep. Labor Income (all observations)



(h) Indep. Labor Income (trimmed sample)



(i) Dep. Labor Income (all observations)



(j) Dep. Labor Income (trimmed sample)

Appendix F: Robustness checks

See Tables 18, 19 and 20.

Table 18 Fixed-effects estimations

	(1)	(2)
Taxable Income	0.0462***	0.0342***
	(0.00807)	(0.00801)
PT Income	− 0.101***	− 0.0794***
	(0.00975)	(0.00938)
14T income	0.158***	0.101***
	(0.0184)	(0.0173)
Withdrawals	0.130***	0.0713***
	(0.0244)	(0.0236)
Dividends	0.115***	0.0948***
	(0.0171)	(0.0166)
Independent Labor Income	0.0839***	0.0781***
	(0.0179)	(0.0175)
Dependent labor income	− 0.0151	− 0.0164
	(0.0136)	(0.0137)
Lagged variables	No	Yes
Observations	78,260	77,982

Estimated coefficient γ from (2) using a fixed-effects model with Taxable Income, PT Income, 14T income, Withdrawals, dividends, Independent Labor Income and dependent labor income as Y_{it}, using the *Trimmed Sample*. Years considered: 2012 (pre-reform) and 2013 (post-reform)
Lagged variables: dependent variable and Taxable Income. Other covariates are omitted since they are time-invariant
Standard errors clustered at the individual level are reported in parentheses
***10%

Table 19 Model in differences

	(1)	(2)
Taxable Income	0.0462***	0.0354***
	(0.00807)	(0.00796)
PT Income	− 0.101***	− 0.0839***
	(0.00975)	(0.00942)
14T income	0.158***	0.101***
	(0.0184)	(0.0173)
Withdrawals	0.130***	0.0755***
	(0.0244)	(0.0235)
Dividends	0.115***	0.0955***
	(0.0171)	(0.0165)
Independent Labor Income	0.0839***	0.0768***
	(0.0179)	(0.0175)
Dependent labor income	− 0.0151	− 0.0161
	(0.0136)	(0.0136)
Lagged variables	No	Yes
Observations	39,130	39,130

Estimated coefficient γ from (2) using a model in differences with Taxable Income, PT Income, 14T income, Withdrawals, dividends, Independent Labor Income and dependent labor income as Y_{it}, using the *Trimmed Sample*. Years considered: 2012 (pre-reform) and 2013 (post-reform)

Lagged variables: dependent variable and Taxable Income. Other covariates are omitted since they are time-invariant

Standard errors clustered at the individual level are reported in parentheses

***10%

Table 20 Difference-in-differences with longer pre-reform period: 2010–2013

	(1)	(2)	(3)
Taxable Income	0.0403***	0.0410***	0.0661***
	(0.00804)	(0.00809)	(0.00910)
PT Income	− 0.101***	− 0.0966***	− 0.157***
	(0.00976)	(0.00976)	(0.0141)
14T income	0.155***	0.156***	0.265***
	(0.0185)	(0.0186)	(0.0278)
Withdrawals	0.122***	0.121***	0.260***
	(0.0246)	(0.0249)	(0.0350)
Dividends	0.112***	0.112***	0.165***
	(0.0173)	(0.0175)	(0.0256)
Independent Labor Income	0.0777***	0.0762***	0.0959***
	(0.0179)	(0.0182)	(0.0259)
Dependent labor income	− 0.0150	− 0.0163	0.0383**
	(0.0137)	(0.0140)	(0.0187)
Additional regressors	No	Yes	Yes
Lagged variables	No	No	Yes
Observations	153,364	152,939	152,130

Estimated coefficient γ from (2) using a model in differences with Taxable Income, PT Income, 14T income, Withdrawals, dividends, Independent Labor Income and dependent labor income as Y_{it}, using the *Trimmed Sample*. Years considered: 2010, 2011, 2012 (pre-reform) and 2013 (post-reform). All specifications include leads with 2012 the omitted time dummy
Additional regressors: sex, age, economic sector fixed effects and town fixed effects
Lagged variables: dependent variable and Taxable Income
Standard errors clustered at the individual level are reported in parentheses
**5%; ***10%

References

Agostini, C. (2013). Una reforma eficiente y equitativa del impuesto al ingreso en chile. In J. Arellano & V. Corbo (Eds.), *Tributación para el Desarrollo: Estudios para la Reforma del Sistema Chileno*. Santiago: Uqbar Editores.

Almunia, M., & Lopez-Rodriguez, D. (2018). Under the radar: The effects of monitoring firms on tax compliance. *American Economic Journal: Economic Policy*, *10*(1), 1–38.

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton: Princeton University Press.

Bird, R. (1974). *Taxing agricultural land in developing countries*. Cambridge: Harvard University Press.

Bird, R. (1992). *Tax policy and economic development*. Baltimore: John Hopkins University Press.

Bird, R., & Wallace, S. (2004). Is it really so hard to tax the hard-to-tax? The context and role of presumptive taxes. In J. Alm, J. Martínez-Vásquez, & S. Wallace (Eds.), *Taxing the hard-to-tax: Lessons from theory and practice*. Bingley: Emerald Group Publishing.

Elschner, C. (2013). Special tax regimes and the choice of organizational form: Evidence from the European tonnage taxes. *Journal of Public Economics*, *97*, 206–216.

Engelschalk, M. (2004). Creating a favorable tax environment for small business development in transition countries. *Contributions to Economic Analysis*, *268*, 275–311.

Fairfield, T., & Jorratt, M. (2015). Top income shares, business profits, and effective tax rates in contemporary Chile. *Review of Income and Wealth*, *62*(1), 120–144.

Imbens, G., & Rubin, D. (2015). *Causal inference for statistics, social, and biomedical sciences. An introduction*. Cambridge: Cambridge University Press.

Imbens, G., & Wooldridge, J. (2008). Recent developments in the econometrics of program evaluation. NBER working papers 14251, National Bureau of Economic Research, Inc.

Jorratt, M. (2009). *La Tributación Directa en Chile: Equidad y Desafíos. Serie macroeconomía para el desarrollo* (Vol. 92). Santiago: CEPAL.

Jorratt, M. (2013). Gasto tributario y evasión tributaria en chile: Evaluación y propuestas. In J. Arellano & V. Corbo (Eds.), *Tributación para el Desarrollo: Estudios para la Reforma del Sistema Chileno*. Santiago: Uqbar Editores.

Kanbur, R., & Keen, M. (2014). Thresholds, informality, and partitions of compliance. *International Tax and Public Finance*, *21*(4), 536–559.

Onji, K. (2009). The response of firms to eligibility thresholds: Evidence from the Japanese value-added tax. *Journal of Public Economics*, *93*(5–6), 766–775.

Saez, E. (2004). Reported incomes and marginal tax rates, 1960–2000: Evidence and policy implications. NBER working papers 10273, National Bureau of Economic Research, Inc.

Saez, E., Slemrod, J., & Giertz, S. H. (2012). The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of Economic Literature*, *50*(1), 3–50.

Saez, E., & Stantcheva, S. (2016). Generalized social marginal welfare weights for optimal tax theory. *American Economic Review*, *106*(1), 24–45.

Slemrod, J. (1989). Complexity, compliance costs, and tax evasion. In J. Roth & J. Scholz (Eds.), *Tax payer compliance: Social science perspectives*. Philadelphia: University of Pennsylvania Press.

Slemrod, J. (2001). A general model of the behavioral response to taxation. *International Tax and Public Finance*, *8*(2), 119–128.

Slemrod, J., & Bakija, J. (2004). *Taxing ourselves: A citizen's guide to the debate over taxes*. Cambridge: The MIT Press.

Slemrod, J., Collins, B., Hoopes, J., Reck, D., & Sebastiani, M. (2017). Does credit-card information reporting improve small-business tax compliance? *Journal of Public Economics*, *149*, 1–19.

Slemrod, J., & Yitzhaki, S. (2002). Tax avoidance, evasion and administration. In A. Auerbach & M. Feldstein (Eds.), *Handbook of public economics* (Vol. 3). Amsterdam: North-Holland.

Terkper, S. (2003). Managing small and medium-sized taxpayers in developing countries. *Tax Notes International*, *29*(2), 211–229.

Thuronyi, V. (2004). Presumptive taxation of the hard-to-tax. *Contributions to Economic Analysis*, *268*, 101–120.

World Bank. (2015). Chile: Efectos Distributivos de la Reforma Tributaria de 2014. Prácticas globales de macroeconomía y gestión fiscal, pobreza y gobierno, en cooperación con el ministerio de hacienda de Chile, World Bank Working Paper. Available at http://documentos.bancomundial.org/curated/es/496131468228282235/pdf/ReformaTributaria-Chile-Final.pdf.