Institut sur la retraite et l'épargne
HEC MONTRÉAL

# Loss-Averse Tax Manipulation and Tax-Preferred Savings

CAHIER DE RECHERCHE No 8
*WORKING PAPER No. 8*

Derek Messacar

Février / *February* 2022

Retirement and Savings Institute

Les opinions et analyses contenues dans les cahiers de recherche de l'Institut ne peuvent en aucun cas être attribuées aux partenaires ni à l'Institut lui-même et elles n'engagent que leurs auteurs.

Opinions and analyses contained in the Institute's working papers cannot be attributed to the Institute or its partners and are the sole responsibility of the authors.

©2022 Statistics Canada / Statistique Canada. Tous droits réservés. All rights reserved. Reproduction partielle permise avec citation du document source, incluant la notice ©. Short sections may be quoted without explicit permission, if full credit, including © notice, is given to the source.

Dépôt légal : Bibliothèque et Archives nationales du Québec et Bibliothèque et Archives Canada, 2022.
ISSN 2561-9039

SDG-13359

**Loss-Averse Tax Manipulation and Tax-Preferred Savings***

Derek Messacar†

January 12, 2022

**Abstract**

Using administrative data from Canada linked to a financial capability survey, I show that tax-deductible savings plans are often used to manipulate final balances owed to the central tax authority during tax season. This finding implies a strong avoidance motive for saving, where tax filers manipulate final balances rather than total tax liabilities, consistent with loss-aversion. The magnitude of this effect is economically significant. For example, each $100 owed increases the likelihood of contributing by about half a percentage point. There is evidence that the behavior is driven by tax filers with low financial literacy who make disproportionately large contributions in the last 60 days before the annual deadline.

*Keywords*: Loss-aversion; tax avoidance; savings; regression kink design.

*JEL Codes*: D14, D91, H26, H31.

* I am grateful to Marc Frenette, Pierre-Carl Michaud, René Morissette, and Alex Rees-Jones for their comments and advice. I also thank Winnie Chan for her work on the data linkage. Statistics Canada acknowledges financial support from the Retirement and Savings Institute at HEC Montréal to produce the data linkage. Any errors are my own.

† Research Analyst, Statistics Canada; Adjunct Professor, Memorial University of Newfoundland; and Research Fellow, Retirement and Savings Institute, HEC Montréal. E-mail address: derek.messacar@statcan.gc.ca. The views and opinions expressed in this paper are those of the author and do not necessarily reflect the views of Statistics Canada or the Government of Canada.

## 1 Introduction

Tax-deductible savings plans encourage people to save by increasing the after-tax rate of return on investments in designated accounts. This is done by allowing contributions to be made with pre-tax income and deferring taxation to the time when funds are withdrawn, typically in retirement when marginal tax rates are otherwise low (Veall 2001). In a life-cycle framework, offering tax advantage is the main policy tool available to governments to stimulate private saving (Feldman 2010). However, savings rates have declined over the past few decades (OECD 2012, 2013), calling into question the effectiveness of tax deductions at influencing consumer behavior; by design, a price incentive only benefits those who are aware of that incentive and understand how to use it (Attanasio, Banks and Wakefield 2004; Chetty et al. 2014a). As governments strive to redesign pension income systems to help people save, a better understanding of how cognitive factors influence savings decisions is increasingly relevant.

The goal of this study is to investigate the relationship between annual tax filing behavior and retirement savings decisions through a behavioral lens. Specifically, I show tax-deductible ("back-loaded") plans are often used by tax filers simply to manipulate their final balances owed to the central tax authority each year to avoid having an outstanding balance.[1]

There is virtually no difference between owing $1 at the end of the year and collecting a refund of $1 since the interest foregone or accumulated on the balance is negligible, so in theory tax filers should not disproportionately favor one outcome over the other. In practice, however, tax filers respond to the perceived loss of owing money at tax season by manipulating total liabilities in order to have a balance close to zero, which likely arises at least in part due to loss-aversion

[1] The back-loaded savings vehicle in Canada is called the registered retirement savings plan (RRSP) and is similar in spirit to the individual retirement account (IRA) in the United States. I adopt the terminology of referring to tax-deductible savings plans by the timing of taxation. These plans are "back-loaded" in the sense that taxation occurs at the end of the use of the plans when funds are eventually withdrawn. Similarly, "front-loaded" plans are plans that offer tax advantage on after-tax income, as in Beshears et al. (2017) and discussed later.

(Rees-Jones 2018). Building on that result, in this paper I show that contributing to back-loaded savings plans is induced by loss-averse tax avoidance. In other words, tax filers exploit the deduction from contributing to a retirement savings plans to push outstanding balances owed at the end of the year down towards zero. This behavior differs from standard models of tax avoidance in which the goal is to minimize liabilities accrued during the full year or life-cycle (Stiglitz 1985; Slemrod and Yitzhaki 2002; Saez, Slemrod and Giertz 2012).

The analysis proceeds in three stages. First, using administrative tax data from Canada covering 20 percent of all tax filers from 2009 to 2016, I estimate the relationship between final tax balances owed to the central tax authority and savings in back-loaded plans. The analysis controls for the fact that final balances depend in turn on the amount saved through the tax deduction by using an "adjusted" measure of balances. Specifically, I simulate marginal tax rates and then predict the final balances that were likely realized *before* any deductions. This approach removes from the measure of final balances the direct effect of back-loaded savings as well the use of any other deductions that may in some way be correlated with savings decisions. This approach is similar in spirit to Engström et al. (2015) who test the probability that Swedish tax payers claim deductions when balances are due, and relates broadly to studies on the elasticity of taxable income that use simulated instruments (Auten and Carroll 1999; Gruber and Saez 2002; Kleven and Schultz 2014).

Through both graphical inspection and regression-based analysis, I show that bunching in final tax balances is in part explained by a change in the use of back-loaded savings plans. The magnitude of this effect is economically significant. For example, each $100 owed increases the likelihood of contributing by about half a percentage point. As a placebo check, I also verify that no comparable effect is observed for a different type of tax-preferred savings vehicle for which contributions are made with after-tax income. The lack of response in this case arises because tax liabilities are determined before the contributions are made ("front-loaded") so this type of savings vehicle does not alter final balances.[2]

Second, I test for heterogeneous responses based on tax filers' ability to understand the relevant tax regulations governing these plans. To this end, I link tax records to a nationally-representative financial capability survey conducted by Canada's central statistical agency in 2014. This survey asks respondents numerous questions to measure financial literacy, such as how inflation affects the future value of savings, interest compounding, diversification and other related issues. The prediction for how financial literacy should impact tax manipulation is ambiguous since it is unclear whether loss-aversion is a mistake or a preference. Since loss-aversion is irrational, one possibility is that informed tax filers are less likely to exhibit such behavior. However, a better understanding of the tax system may lead to greater tax manipulation if loss-aversion is common. The empirical findings are consistent with the former hypothesis, which suggests that a tax mistake is occurring.

During the first few months each year, individuals begin receiving their statements of remuneration paid from employers and are typically able to start preparing their previous year's income tax statements. A feature of the Canadian income tax system is that, during the first 60 days of the calendar year, tax filers can contribute to back-loaded savings plans and then apply those contributions as a deduction on the previous year's income to reduce how much they will owe (or increase how much they receive as a refund) immediately upon filing their previous year's income tax statements. As a result, tax filers have a direct mechanism for accurately estimating their tax liabilities and then manipulating final balances. The deadline to contribute each year is

[2] The front-loaded savings vehicle in Canada is called the tax-free savings account (TFSA) and is similar in spirit to the Roth IRA in the United States.

called the "RRSP deadline." It is a salient feature of the tax code that often receives media attention as the date approaches to remind Canadians to contribute if they have not already done so.[3,4,5,6,7] Financial institutions often provide "RRSP loans" to help tax filers contribute more than they have available to save to facilitate tax planning.

Fortunately, the central tax authority collects information *separately* on the amount saved in back-loaded plans in the last 10 months of the same calendar year and the first 60 days of the next calendar year. These files are stand-alone administrative records that I incorporate into the linked dataset. I find that "last-minute" saving is very common, suggesting that tax-deductible saving occurs sporadically within a calendar year. I also show that the majority of loss-averse tax manipulation occurs in the 60 days before the deadline.

In the third stage of the analysis, I exploit the longitudinal design of the tax data to consider whether savings induced by loss-aversion remain in these plans or are withdrawn the following year. Specifically, I estimate the effects of contributing during the last 10 months of calendar year $t-1$ or the first 60 days of calendar year $t$ (both of which affect tax liabilities for tax year $t-1$) on the probability of withdrawing in year $t$. I control for observed characteristics that vary across individuals and over time that may affect such behavior in a regression framework. I find that individuals who contributed during the 10 month interval are also more likely to subsequently withdraw. This finding is perhaps not surprising since a prerequisite for withdrawing is having funds in the account. However, I find no evidence that contributing in the last 60 days is associated with an increase in withdrawals, which leads me to conclude that loss-averse tax manipulation likely boosts total wealth accumulation.

---

[3] https://www.moneysense.ca/save/investing/rrsp/the-procrastinators-guide-to-the-rrsp-deadline.
[4] https://globalnews.ca/news/7647821/rrsp-deadline-2020-taxes.
[5] https://vancouver.citynews.ca/2021/02/26/canadians-retirement-rrsp-deadline/.
[6] https://ca.finance.yahoo.com/news/why-missing-todays-rrsp-deadline-isnt-the-end-of-the-world-193730686.html.
[7] https://www.theglobeandmail.com/investing/personal-finance/taxes/article-use-an-rrsp-to-reduce-tax-on-government-benefits.

The paper proceeds as follows. The next section is a brief survey of the related literature. Then, Section 3 summarizes the institutional setting, focusing on relevant aspects of the Canadian tax and transfer and retirement income systems. Section 4 outlines the empirical methodology and key identifying assumptions. Section 5 describes the datasets used. Sections 6, 7 and 8 present the main findings, heterogeneity tests, and analysis of next-period withdrawals, respectively. Lastly, Section 9 concludes.

## 2 Related Literature

Understanding how and why tax filers manipulate and respond to their final income tax balances is particularly interesting from a behavioral economics perspective. The time between paying taxes through source deductions and paying final balances or collecting refunds at the end of the tax season is small, and interest accruals on outstanding balances are negligible so there is little incentive to manipulate this outcome from a life-cycle perspective. However, Engström et al. (2015) and Rees-Jones (2018) have shown that the marginal return to a dollar is different for tax filers who have positive versus negative balances, a finding that is at least in part explained by the psychological occurrence of loss-aversion. Another possible explanation for tax filers preferring to have a negative balance is forced saving (Feldman 2010).

This study contributes to a growing literature in public finance that shows tax payers often make costly "mistakes" for behavioral reasons—i.e., choices that differ from expectations if we assume tax filers are rational and well-informed. For example, Chetty, Looney and Kroft (2009) and Finkelstein (2009) show that consumers underreact to taxes that are not salient. Feldman, Katuščák and Kawano (2016) find tax filers misperceive changes in marginal effective tax rates resulting from predictable changes in tax credits over time. Søgaard (2019) finds optimization frictions in the Danish student labor market resulting from inattention. Such behaviors may arise because tax systems are complex, in which a myriad of programs can make it difficult to know one's own marginal effective tax rate (Liebman and Zeckhauser 2004; Kotlikoff and Rapson 2007; Milligan 2009; Ito 2014; Rees-Jones and Taubinsky 2019).

There is also a large literature on the effects of tax-preferred savings plans on total wealth accumulation. Earlier studies address whether tax deductions induce greater savings through a price effect or reduce total savings through an income effect (Bernheim (2002) provides a comprehensive survey). In an early analysis, Feenberg and Skinner (1989) show that contributions to individual retirement accounts are more likely when a tax filer owes a positive balance to the Internal Revenue Service. A few studies also estimate whether front-loading taxation increases total savings, which relates broadly to the importance of the timing of taxation (Beshears et al. 2017; Berger, Farrar and Zhang 2019; Lavecchia 2019).

This study contributes to a growing field of research on the behavioral determinants of retirement savings decisions. Chetty et al. (2014a) assess the relative effectiveness of tax incentives versus automatic contributions at boosting savings rates, finding that most savers are "passive" to price incentives. Messacar (2018a) finds that passive behavior is amenable to change through education. Feldman (2010) and Card and Ransom (2011) find individuals have different propensities to save out of income from different sources, consistent with mental accounting. Beshears et al. (2013) show that simplifying complex investment decisions improves financial outcomes, likely due to bounded rationality. Choi, Laibson and Madrian (2011) observe employees with dominated contribution rates to employer-sponsored pension plans and find that educating them about this "free lunch" raises contributions, suggesting a lack of financial literacy or program knowledge. Taken together, these studies emphasize that there are many determinants of saving that fall outside of the stylized life-cycle model. The fact that loss-averse tax manipulation helps drive retirement saving is another novel example to add to this list.

## 3 Institutional Setting

In Canada, the unit of taxation is the individual. Personal income tax is based on a measure of taxable income minus permitted deductions, then credits are applied to determine the net amount payable. There are two levels of taxation—federal and provincial/territorial—where each applies its own tax schedule to a common measure of taxable income (i.e., a common base) and offers its own credits in order to determine total tax liabilities.

Employers are required to withhold income tax and several other source deductions, notably mandatory contributions to a federal unemployment insurance program and means-tested public pension, as well as other payments such as union dues and employer-sponsored pension plan contributions, if applicable. Tax withholding follows a standard employers' guide for payroll deductions and remittances. When a tax filer only works for a single employer, does not have other allowances, credits or deductions that influence taxable income, and all earnings are paid out in the same province/territory as the location of residence at the end of the calendar year (which occurs in most cases), it follows that the guide for payroll deductions likely forecasts tax liabilities with high precision. As will be shown, there is a large mass of tax filers in Canada with balances that are very close to zero, even among those who have positive taxable incomes and tax withholding, compared to the mass observed at the same threshold in the United States (Rees-Jones 2018) where the unit of taxation is the couple and total tax liabilities are harder to forecast precisely by employers.

A common type of tax deduction is the net amount contributed to back-loaded plans in the year, called registered retirement savings plans (RRSPs). Similar to individual retirement accounts (IRAs) in the United States, these are defined-contribution plans that individuals set up and maintain through their banks or financial institutions. Private retirement savings vehicles that offer favorable tax treatment constitute one of Canada's three main tiers of the retirement income system; the other two are a basic demogrant for individuals who are at least 65 years old and a means-tested public pension from which benefits depend on work/earnings histories and for which eligibility begins at 60 years of age.[8]

A unique feature of RRSPs is that, while they are labelled as a retirement savings vehicle, the tax code does not explicitly discourage pre-retirement withdrawals.[9] Evidence suggests these plans are often used for precautionary savings and income-smoothing purposes (Mawani and Paquette 2011). This fact is particularly relevant to this study because it implies, in principle, that back-loaded plans can be used to manipulate final tax balances and then funds can be withdrawn early into the next tax season with very little impact on disposable income throughout the time period.[10] Because of the tax revenue implications, annual contributions are limited to the lesser of a fraction of earnings and a pre-specified threshold (e.g., 18 percent of earned income or $25,370 in 2016). Since 1991, however, unused contribution room has carried forward indefinitely such that the limit typically only binds for a small subset of contributors.

[8] Employer-sponsored pension plans and deferred profit-sharing plans constitute another prevalent type of private tax-preferred savings vehicle, similar to 401(k) plans in the United States.
[9] This contrasts with IRAs, which impose a 10-percent penalty on distributions before 59.5 years of age.
[10] A notable exception is that financial institutions must withhold a fraction of pre-retirement withdrawals and remit the funds to the central tax authority as a partial payment of the income tax that will eventually be owed. In most provinces/territories, the rates ranges from 10 percent on distributions below $5,000 to 30 percent on those above $15,000. If the tax withholding rate exceeds the final marginal tax rate, then a refund is issued. This process is analogous to employers withholding a fraction of wages and salaries. Messacar (2018b) shows tax withholding can serve as a de facto savings commitment device.

The tax-preferred savings vehicle for which the timing of taxation is front-loaded is called the tax-free savings account (TFSA). This plan is similar to Roth IRAs in the United States. While contributions are made with after-tax income, capital gains on assets held in TFSAs accrue tax-free and distributions are non-taxable. The TFSA was introduced in 2009, which corresponds to the first year of observation used in this study for that reason. Annual contribution limits have changed from year to year but they have always ranged from \$5,000 to \$10,000. However, as with RRSPs, unused contribution room carries forward indefinitely. Since TFSAs are available universally, there is little convincing quasi-experimental evidence on the extent to which these plans crowd out savings in other accounts at the aggregate (population) level due to the difficulty of defining a credible control group. However, several studies that provide some insight into this issue find TFSAs at least partially displace other savings in taxable accounts or RRSPs (Messacar 2017; Berger, Farrar and Zhang 2019; Lavecchia 2019). TFSAs were not introduced exclusively as a retirement savings vehicle nor are they advertised as such.

## 4 Empirical Approach

The goal is to estimate how tax-preferred savings, $S_{it}$, for individual $i$ at time $t$ responds to a change in the final tax balance owed to the central tax authority, $B_{it}$, as this moves from being negative (i.e., refund) to positive (i.e., payment due). However, directly estimating this relationship is confounded by the fact that contributions to back-loaded savings plans, $S_{it}^{BL}$, reduce balances owed for most savers through a tax deduction, $B_{it} = f(S_{it}^{BL}, \epsilon_{it})$.

The standard approach for overcoming this reverse-causality problem is to exploit policy-induced changes in tax withholding rates over time, as in related studies from the United States (Feldman 2010; Jones 2012). Unfortunately, to my knowledge, there have been no structural changes in the tax withholding schedules in Canada that would facilitate such an analysis. To address this concern, I construct a simulated measure of balances by predicting how much tax filers owed before any deductions were applied. To this end, I use the Canadian Tax and Credit Simulator (CTaCS) created by Milligan (2016) and exploit the wide set of information in administrative records to simulate tax filers' effective marginal tax rates, $\tau_{it}$.[11] This is an "effective" rate in the sense that it accounts for both statutory taxes and some federal and provincial/territorial allowances, credits, deductions and transfers. Table A1 in the Appendix shows that CTaCS performs well, on balance, at predicting tax liabilities.[12]

The predicted final tax balance is given by $\tilde{B}_{it} = B_{it} + (\tau_{it} \times D_{it})$, where $D_{it}$ is total deductions (i.e., total pre-tax income minus net income). By construction, $D_{it} = S_{it}^{BL} + \xi_{it}$, where $\xi_{it}$ comprises all deductions except for the amount of back-loaded savings. The expectation is that final tax balances are distributed smoothly at zero and, as a result, $S_{it}^{BL}$ is a smooth function of $\tilde{B}_{it}$ (and $B_{it}$) since no tax avoidance mechanisms are being used to manipulate balances. However, if savers are loss-averse, then $S_{it}^{BL}$ should be larger when $\tilde{B}_{it} > 0$ in order to push balances down. No equivalent pattern is expected for savings in front-loaded plans, $S_{it}^{FL}$, because they do not facilitate tax avoidance at the time that the decision to contribute is being made.

I begin by taking these predictions to the data through graphical inspections. I plot the distribution of actual final tax balances, $B_{it}$, to assess whether Canadians appear loss-averse as in Rees-Jones (2018), and check whether this behavior occurs through the use of deductions. Then I

[11] CTaCS is similar in spirit to the NBER's TAXSIM program.

[12] Total tax liabilities are observed in the administrative records. However, the dataset does not contain a variable for the marginal (statutory or effective) tax rate. For this reason, CTaCS is used to simulate tax rates. The predicted taxes typically exceed actual taxes by a small amount, which likely occurs because not all deductions available to tax filers are being included in CTaCS. However, the basic relationships across groups regarding who pays less or more in personal income taxes is very consistent between the actual and predicted variables.

plot savings outcomes against the adjusted final tax balances in order to view whether a change in savings occurs at the zero balances threshold.

I use three different measures of savings outcomes throughout the analysis. The first is savings rates in back-loaded and front-loaded plans, $s_{it}^{BL} = S_{it}^{BL}/Y_{it}$ and $s_{it}^{FL} = S_{it}^{FL}/(Y_{it} - T_{it})$, respectively, where $Y_{it}$ is before-tax income and $T_{it}$ is the total tax liability. Contributions to back-loaded plans are made with pre-tax income while contributions to front-loaded plans are made with after-tax income, so the rates are constructed accordingly. The second measure is an indicator for having any positive savings, and the third is the log of savings.[13] Unless stated otherwise, I always use *net* savings defined as the difference between total contributions less withdrawals.

To provide formal structure to this analysis, I also estimate how savings changes at the zero-balance threshold in a regression kink design (RKD). Denote $\widetilde{K}_{it} = 1(\tilde{B}_{it} > 0)$ where $1(\cdot)$ is the indicator function. The statistical model is:

$$s_{it} = \alpha + g(\tilde{B}_{it}) + \beta(\tilde{B}_{it} \times \widetilde{K}_{it}) + X_{it}'\theta + \pi_t + \rho_p + \varphi_{it} \qquad (1).$$

This analysis is repeated for each $s_{it} \in \{s_{it}^{BL}, s_{it}^{FL}\}$.

As will be shown, this functional form is well-supported by the graphical analysis. The preferred specification assumes $g(\cdot)$ is a quadratic function; Gelman and Imbens (2019) recommend using only local linear or quadratic fits in regression discontinuity and RKD settings, as higher-order polynomials often absorb much of the variation of interest. I express $\tilde{B}_{it}$ in units of \$100 so that the resulting coefficient estimates can be interpreted as the change in saving per \$100 owed at tax season.

[13] Since the dependent variable is net saving, I rule out cases where individuals may have very small amounts of net saving that have virtually no effect on tax liabilities by setting the indicator variable to be "1" if saving is \$100 or more and "0" otherwise. In addition, since net savings may be zero or negative, I use the inverse hyperbolic sine (IHS) transformation to estimate the effect of final balances on the log of savings. The IHS transformation provides approximately the same interpretation as the log transformation but it is well-defined at zero and negative values.

I condition the analysis on tax filers who have final balances between −\$2,000 and \$2,000. This RKD "bandwidth" is larger than the one used by Rees-Jones (2018) to analyze loss-aversion. I use a wider interval because I am applying equation (1) to a transformed measure of final balances. Individuals who may have initially been very close to the zero-balances threshold but who have a large amount of deductions may be shifted relatively far away from this threshold after the transformation, so a wider bandwidth is needed to capture the full extent of the relationship between tax balances and savings.

The control variables, $X_{it}$, include fixed effects for year, age, gender, province, and marital status. I include a quadratic polynomial in total income and a variable for the marginal tax rate to absorb any variation across individuals and over time in earnings and tax-based savings incentives that may confound the results. I also include the amount saved each year in employer-sponsored pension plans. In Canada, the total amount that a person can contribute to back-loaded savings plans is reduced dollar-for-dollar by the amount contributed to employer-sponsored plans, so including this variable helps to control for this institutional savings constraint. Exploiting the panel design of the data, I specify individual fixed effects in some specifications. In this case, the effect of interest ($\beta$) is identified by changes in saving that occur as individuals move around the zero-balance threshold over time. Lastly, the terms $\pi_t$ and $\rho_p$ are year fixed effects and province fixed effects, respectively, and $\varphi_{it}$ is the residual.

Two limitations of this approach are important to note. The first is that final tax balances, $B_{it}$, are known to be manipulable, which violates the key identifying assumption of the RKD that observations are distributed randomly around the discontinuity. However, this issue is addressed by: (i) the way that $\tilde{B}_{it}$ is constructed to directly offset tax-based sorting; (ii) controlling in a regression-based framework for other characteristics that vary near the discontinuity; (iii) including individual fixed effects to assess how the results change after removing individual-specific confounders; and (iv) the placebo check using front-loaded savings plans. The second limitation is that, because $\tilde{B}_{it}$ is a predicted variable given that $\tau_{it}$ is simulated, there is a possibility of measurement error such that equation (1) is a "fuzzy" RKD. As will be shown graphically, this error appears negligible.

## 5 Data and Sample Selection

This study is based primarily on an analysis of the Longitudinal Administrative Databank (LAD) from Statistics Canada ("tax dataset"). The LAD is a 20-percent sample of personal income tax records spanning 1982 to 2018 at the time of writing. This dataset derives from the central tax authority, the Canada Revenue Agency, and contains a wide set of information about tax filers' demographics, income, credits, deductions, taxes and transfers. Contributions and withdrawals to back-loaded and front-loaded plans are included in the data, although the latter is only observed from 2009 onward since this savings vehicle did not exist prior to that time. It also includes the total amount of income tax deducted at source (i.e., tax withholding) and the final income tax balances owed to or refunded by the central tax authority.

A limitation of the tax dataset is that it does not include a variable for tax filers' marginal tax rates needed to calculate adjusted final balances. To overcome this issue, I use the Canadian Tax and Credit Simulator (CTaCS) created by Milligan (2016). The latest version available for the simulator is 2016 at the time of writing so, for this reason, 2016 is the last year of observation used. Coupled with the fact that front-loaded plans were introduced in 2009, I restrict the period of analysis in this study to be from 2009 to 2016.

In addition, I impose three other sample restrictions. First, I focus on tax filers who are 20 to 54 year olds. The lower bound drops teens who may still be in school and are unlikely to use tax-preferred savings plans and the upper bound excludes those whose savings decisions may be affected by the fact that they are approaching the age of retirement eligibility. The earliest ages at which individuals may collect public and employer-sponsored pensions are typically 60 and 55 years old, respectively. Second, following Rees-Jones (2018), I drop tax filers with no taxable income or no tax withholding to rule out cases where final balances are zero simply because the individuals face zero tax liabilities. Third, for methodological reasons, I exclude tax filers who have no deductions. As will be shown, this restriction leads to conservative estimates of the treatment effect.

While the tax dataset provides information on many outcomes of interest for tax-based research, it does not include variables that do not appear on tax forms but that are nevertheless of interest for analytical purposes, notably financial literacy. To close this gap, I link a nationally-representative survey called the Canadian Financial Capability Survey (CFCS), conducted by Canada's central statistics agency in 2014, to respondents' personal income tax records. This linked file is called the Financial Capability, Employment and Income Database (henceforth the "linked dataset").[14]

The CFCS includes 14 questions that measure respondents' financial knowledge. The quiz includes standard questions related to the effects of inflation and interest compounding, as well as others about risks associated with stock market participation, different types of investment portfolios, and benefits of diversification.[15] Low versus high financial literacy is defined based on

[14] More accurately, I coordinated the data linkage that was carried out by data development experts and methodologists at Statistics Canada using standard procedures for matching respondents deterministically to their administrative tax records based on such information as Social Insurance Number (SIN), name, address, year of birth, or gender.

[15] Mullock and Turcotte (2012) provide a detailed assessment of responses to this questionnaire.

whether the respondent scores between zero and nine ("low") or from 10 to 14 ("high"), which is consistent with the approach used by Laurin et al. (2021). A limitation of the linked dataset is that savings in front-loaded plans are not observed, which means the analysis of heterogeneity by level of financial literacy only considers back-loaded plans.[16] I impose the same set of sample restrictions for this linked dataset as those used with the tax dataset.

Table 1 presents descriptive statistics for both datasets. On balance, individuals in this sample are about 37 years old, half of them are female and the other half are male, and about 57 percent are married (including common-law relationships). These tax filers contribute to back-loaded savings plans about one-third of the time and contribute to front-loaded plans 16 percent of the time. Across most variables including the tax liabilities and marginal effective tax rates, the values are very similar between the tax and linked datasets.

## 6 Primary Results

In Figure 1, I begin by plotting the distribution of final tax balances among the full sample of tax filers, from −\$2,000 to \$2,000 in bins of width \$100. Tax filers appear to bunch at or near the zero-balances threshold and have a higher propensity to receive a refund than to owe money (panel A). The large amount of excess mass at the threshold is suggestive of very accurate tax planning during the year and/or frequent tax manipulation. The adjusted measure of final balances shows a similar pattern, with excess mass at the zero-balances threshold, but the magnitude of bunching is much smaller (panel B). This suggest tax deductions are widely used to manipulate balances but there remains some fraction of tax filers in the full sample who have zero balances

[16] TFSA information does not appear on T1 tax records but is reported to the Canada Revenue Agency by financial institutions. Through a linkage procedure, Statistics Canada merged TFSA information obtained from the Canada Revenue Agency with the LAD. The CFCS was linked to the T1 Personal Master File, which is approximately a 100 percent sample of Canadian tax files. Linking to the population file maximizes the match rate but unfortunately it means TFSA variables are not available.

for other reasons. Restricting the sample to tax filers with positive deductions largely eliminates this excess mass and provides a smooth distribution of tax filers around the zero-balance threshold needed for implementing the RKD. Thus, I focus the analysis on this restricted sample but also consider how results vary in the full sample.

To observe whether back-loaded plans are used to manipulate final tax balances, Figure 2 plots the relationships between savings rates and final tax balances using the actual and adjusted measures in panels A and B, respectively. Each dot corresponds to the average value of the savings rate within a bin of width $100 corresponding to the same bins shown in Figure 1. In panel A, the savings rate for back-loaded plans slopes steeply downward suggesting that tax filers with higher savings have lower final balances, which is a mechanical effect of how this deduction affects tax liabilities. In contrast, the savings rate in front-loaded plans is much flatter. Both trends show a spike at the zero-balance threshold due to the change in sample composition at this point, as shown in Figure 1.

In panel B of Figure 2, the savings profile for back-loaded plans changes significantly when it is expressed relative to the adjusted final tax balances. In this case, the savings profile is relatively flat for negative balances but tax filers who are predicted to owe money at the end of the year have higher net savings rates in these plans compared to tax filers with negative balances. This savings rate increases with the amount owed, suggesting tax filers are contributing more to push balances down to zero as they move further away from this threshold. The savings rate in front-loaded plans does not exhibit a similar pattern, hence the effect is not likely driven by a change in sample composition.

I explore these findings in more detail in Figure 3 by focusing on an indicator for having positive savings (panel A) and the log of savings (panel B) as the outcomes. The results indicate, for back-loaded savings plans, that these measures of saving largely mirror the trends for savings rates. Second, except for minor variations farther away from the zero-balances threshold, the savings profiles for front-loaded plans are relatively flat.[17]

To provide formal structure to this analysis, I estimate the savings profiles for back-loaded and front-loaded plans using the statistical model in equation (1). Doing so permits me to control parametrically for changes in observed characteristics that may change with the forcing variable that could bias the estimates. The results of this analysis for back-loaded and front-loaded plans are shown in Tables 2 and 3, respectively. Both tables consider the results separately using as the dependent variable the savings rate (panel A), the probability of saving (panel B) and the log of savings (panel C). I also vary the set of controls across columns. In particular, column 1 controls linearly for the forcing variable whereas the remaining columns control for the forcing variable using a quadratic function; column 3 includes control variables; and column 4 includes individual fixed effects. For compactness, I only report estimates of the changes in slope. Standard errors are clustered by the forcing variable in bins of width $100.

The results of this analysis are generally consistent with the graphical evidence and do not change meaningfully with the addition of controls in most cases. For example, the net savings rate in back-loaded plans increases by roughly half a percentage point per $100 owed at the zero-balances threshold. For front-loaded plans, no discernible effect is observed after the quadratic term is included, which absorbs the slight non-linearity observed in Figures 2 and 3.

In both of the preceding tables, the RKD estimates are smaller when individual fixed effects are included in the model; however, the results consistently indicate that back-loaded savings plans

[17] Figure A1 shows how the results look using the full sample. In this case, there is a large spike in savings at the zero-balances threshold. The results are fairly similar to Figures 2 and 3 except that the magnitude of the changes for back-loaded savings plans are larger and there appears to be a small upward trend for front-loaded savings plans. The full sample is biased towards finding a positive effect owing to the change in sample composition around the zero-balances threshold shown in Figure 1.

facilitate loss-averse tax manipulation. One explanation for the muted effect size when individual fixed effects are included is that the model relies on tax filers who move around the discontinuity over time for identification. As a result, those who (by chance) consistently land on one side or the other of the threshold and whose savings decisions reflect this fact are omitted. If savings responses occur gradually due to inertia or procrastination (Jones 2012)—especially along the extensive margin where there are fixed costs of setting up savings accounts or making contributions—relying only on tax filers who move around the discontinuity may under-estimate the true magnitude of loss-aversion in the population.

## 7 Heterogeneity

In this section, I rely on the linked dataset to test for differences in savings responses by financial literacy. Although this analysis is somewhat limited by sample size, I find that the results are comparable to those obtained from the tax dataset.

In Figure 4 plots the distributions of adjusted final tax balances separately for those with low versus high financial literacy. The results are consistent with the full-sample distribution shown previously in Panel C of Figure 1. Importantly, there does not appear to be any relevant difference in the distribution of tax filers around the zero-balances threshold that may bias the heterogeneity analysis in some way.

Figure 5 plots the relationship between saving and adjusted final tax balances separately by financial literacy. The patterns are generally consistent with those observed in Figures 2 and 3 although much noisier due to the significant reduction in sample size. It also appears that changes at the zero-balance threshold are slightly larger for tax filers with low financial literacy relative to their counterparts with high literacy. Particularly in panels B and C, there is a discernible gap between low and high literacy everywhere to the left of the threshold that gradually closes as the balances become positive and increasingly large.

Due to the small sample size in the linked dataset, I rely on the RKD estimates to determine whether the differences between tax filers with low and high financial literacy are statistically relevant. In doing so, this also allows me to control parametrically for other differences in personal characteristics that may affect savings.

In Table 4, I report the RKD results for the savings rate (panel A), the probability of saving (panel B) and the log of savings (panel C). I carry out the analysis for the full sample (columns 1 and 2) in order to compare how the results vary between the tax and linked datasets. I also carry out the analysis separately for tax filers with low (columns 3 and 4) and high (columns 5 and 6) financial literacy. I employ the preferred model specification with a quadratic term for the forcing variable and full set of controls, and report the results separately with and without individual fixed effects as shown in the column headings.

The results of this analysis are generally consistent with expectations based on the findings from the previous section. In the full sample, the change in savings rate at the zero-balance threshold is similar to that obtained in the tax dataset. For the other two outcomes, the estimated effects are a bit larger in magnitude but continue to suggest that owing a positive balance at tax season is associated with an increase in the likelihood of using back-loaded savings plans. In addition, the heterogeneity analysis in columns 3 to 6 suggest this participation response is primarily driven by those with low literacy.

Next, I decompose the results based on the amounts contributed in the 10 month and 60 day intervals, as shown in Figure 6 and Table 5.[18] In Figure 6, I plot the likelihood of saving and

[18] It is not possible to decompose the timing of withdrawals from back-loaded savings plans into these 10 month and 60 day intervals. As a result, I use gross (rather than net) contributions for this part of the analysis.

the average amount saved (conditional on making a positive contribution) based on the timing of this behavior. I find that tax filers are about equally likely to contribute in a back-loaded savings plan in the two time periods. The amount saved in the 10 month interval is about $3,000 in 2016, while the amount saved in the 60 day interval is about $2,000 in 2016. This result is interesting since it shows tax filers are observed making about 40 percent of their total annual savings (i.e., $2,000 divided by $5,000) in the last 17 percent of the year (i.e., 2 months divided by 12 months).

The results in Table 5 show that preceding analysis masks important heterogeneity based on the timing of saving. In particular, savings made during the 10 month interval do not appear to respond very much to owing money at tax season. However, for tax filers with low literacy, savings made during the 60 day interval are very responsive to owing money rather than receiving a refund. For example, I estimate that each additional $100 owed increases the likelihood of contributing to a back-loaded savings plan in the 60 day interval by up to a full 2 percentage points depending on the model specification.

Taken together, these findings have two implications. First, loss-averse tax manipulation is likely a mistake since increased financial literacy mitigates this behavior. Second, tax filers are likely being nudged to save myopically due to the design of tax regulations governing back-loaded savings plans in Canada.

## 8 Withdrawals

I now consider whether savings to back-loaded plans induced by loss-aversion are later withdrawn or remain in place leading to greater wealth accumulation. Define $S_{it}^{10M}$ and $S_{it}^{60D}$ as the amount saved in back-loaded plans during the 10 month and 60 day intervals, respectively, for individual $i$ in the *tax* year $t$. Further, define $W_{it}^{BL}$ as a measure of withdrawals from back-loaded plans. I estimate the following statistical model:

$$W_{i,t+1}^{BL} = \mu + \phi 1(S_{it}^{10M} > 0) + \psi 1(S_{it}^{60D} > 0) + X_{it}'\theta + \pi_t + \rho_p + \varphi_{it} \quad (2).$$

The terms $1(S_{it}^{10M} > 0)$ and $1(S_{it}^{60D} > 0)$ are indicators for having positive saving in the two respective time periods.[19] Thus, the coefficients of interest, $\phi$ and $\psi$, measure the increase in the likelihood of withdrawing in tax year $t + 1$ conditional on making a positive contribution in the 10 month and 60 day interval leading up to the deadline, respectively. Both of these coefficients are evaluated relative to tax filers who did not contribute at all in tax year $t$. The same set of control variables used in equation (1) are also included in this model except for final balances. Standard errors are clustered at the individual level.

The results from equation (2) are presented in Table 6. Consistent with the preceding analysis, I report estimates from the full sample and by level of financial literacy,  and I consider how the results vary by model specification with versus without individual fixed effects. I find that individuals who contributed during the 10 month interval are between 2.5 and 6.5 percent more likely to withdraw the next year depending on the model specification, and this effect is driven by individuals with low financial literacy. This finding is consistent with Laurin et al. (2021) who shows that financial literacy affects the responsiveness of savings withdrawals to tax incentives in Canada. However, I find no evidence that contributing in the 60 days leading up to the deadine is associated with a change in the likelihood of withdrawing. Hence, while financal literacy matters for wealth accumulation, I conclude that loss-averse saving among those with low literacy likely increases the amount saved over a lifetime.

[19] As in the preceding analysis, in practice I set the threshold to be $100 or more in savings. This means individuals with very low levels of saving and for whom the tax implications of contributing or withdrawing are negligible are set to zero.

## 9 Conclusion

This paper offers compelling evidence that many tax filers use back-loaded savings plans to reduce their current-year final balances owed to the central tax authority to be close to zero, which results in bunching in the distribution of final balances around zero. In a standard model, tax filers should not have a strong preference between owing money at tax season and collecting a refund since both outcomes result in the same amount of total tax liabilities paid. However, in practice, tax filers manipulate their final balances to avoid owing money in part due to loss-aversion. Therefore, the results of this study imply loss-aversion is a relevant behavioral factor underpinning how and why people contribute to retirement savings vehicles.

This finding has implications for economic theory and policy. Two growing literatures in public finance find tax filers make costly "mistakes" understanding and responding to the tax code as well as saving optimally for retirement. The fact that savings vehicles support suboptimal tax manipulation highlights the interconnected nature of these two issues. While price incentives to save are typically designed with a life-cycle optimization framework in mind, this study suggest programs designed using a behavioral lens may be more effective.

The magnitude of the participation response is large and raises questions about the extent to which loss aversion is the only factor influencing such behavior. In practice, it is likely that institutional factors are nudging tax filers to use their back-loaded savings plans this way. In Canada, contributions made in the first 60 days of the calendar year can be used as deductions for the *previous* year's taxable income. This means individuals can make top-up contributions after they receive their tax forms from employers and compute their total liabilities, allowing them to precisely determine how much they need to contribute to drive down final balances to zero. The analysis presented in this study suggests such behavior is common. This likely occurs because financial institutions and the media advertise and remind Canadians about the annual savings deadline and to contribute before it's too late.

In addition, back-loaded plans are labelled as a retirement savings vehicle in Canada but they do not impose explicit penalties on pre-retirement lump-sum withdrawals and are often used for precautionary saving and income-smoothing. Thus, tax filers' desire not to owe money at tax season combined with their ability to withdraw funds in the short term if needed likely explains why back-loaded plans are widely-used in Canada. These considerations offer insights into how tax-preferred savings vehicles can be designed in other countries to increase usage while recognizing that some savers are behavioral.

Lastly, I find that loss-averse saving occurs primarily among tax filers with low financial literacy, suggesting that a tax "mistake" is indeed occurring. However, the fact that this saving is *not* typically withdrawn the next year implies an increase in total savings. These two findings together make it unclear whether there is a role for paternalistic intervention since this suboptimal savings behavior among tax filers with low financial literacy may still lead to greater lifetime wealth accumulation.

## References


Auten, G., and R. Carroll. 1999. "The effect of income taxes on household income." *Review of Economics and Statistics* 81 (4): 681–693.

Attanasio, O.P., J. Banks, and M. Wakefield. 2004. "Effectiveness of tax incentives to boost (retirement) saving: Theoretical motivation and empirical evidence." *OECD Economic Studies* 39 (2): 145–172.

Berger, L., J. Farrar, and Lu Zhang. 2019. "An empirical analysis of the displacement effect of TFSAs on RRSPs." *Canadian Tax Journal* 67 (2): 309–333.

Bernheim, B.D. 2002. "Taxation and saving." In *Handbook of Public Economics, Volume 3*, ed. A.J. Auerbach and M. Feldstein, p. 1173–1249. Elsevier.

Beshears, J., J.J. Choi, D. Laibson, and B.C. Madrian. 2013. "Simplification and saving." *Journal of Economic Behavior & Organization* 95 (C): 130–145.

Beshears, J., J.J. Choi, D. Laibson, and B.C. Madrian. 2017. "Does front-loading taxation increase savings? Evidence from Roth 401(k) introductions." *Journal of Public Economics* 151: 84–95.

Card, D., and M. Ransom. 2011. "Pension plan characteristics and framing effects in employee savings behavior." *Review of Economics and Statistics* 93 (1): 228–243.

Chetty, R., J.N. Friedman, S. Leth-Petersen, T. Nielsen, and T. Olsen. 2014a. "Active vs. passive decisions and crowdout in retirement savings accounts: Evidence from Denmark." *Quarterly Journal of Economics* 129 (3): 1141–1219.

Chetty, R., J.N. Friedman, T. Olsen, and L. Pistaferri. 2011. "Adjustment costs, firm responses and micro vs. macro labor supply elasticities: Evidence from Danish tax records." *Quarterly Journal of Economics* 126 (2): 749–804.

Chetty, R., J.N. Friedman, and J.E. Rockoff. 2014b. "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood." *American Economic Review* 104 (9): 2633–2679.

Chetty, R., A. Looney, and K. Kroft. 2009. "Salience and taxation: Theory and evidence." *American Economic Review* 99 (4): 1145–1177.

Choi, J.J., D. Laibson, and B.C. Madrian. 2011. "$100 bills on the sidewalk: Suboptimal investment in 401(k) plans." *Review of Economics and Statistics* 93 (3): 748–763.

Engström, P., K. Nordblom, H. Ohlsson, and A. Perrson. 2015. "Tax compliance and loss aversion." *American Economic Journal: Economic Policy* 7 (4): 132–164.

Feenberg, D., and J. Skinner. 1989. "Sources of IRA Saving." In (ed.) L.H. Summer, *Tax Policy and the Economy*, vol. 3, 25–46. Cambridge: MIT Press.

Feldman, N. 2010. "Mental accounting effects of income tax shifting." *Review of Economics and Statistics* 92 (1): 70–86.

Feldman, N., P. Katuščák, and L. Kawano. 2016. "Taxpayer confusion: Evidence from the Child Tax Credit." *American Economic Review* 106 (3): 807–835.

Finkelstein, A. 2009. "E-ZTax: Tax salience and tax rates." *Quarterly Journal of Economics* 124 (3): 969–1010.

Gelman, A., and G. Imbens. 2019. "Why high-order polynomials should not be used in regression discontinuity designs." *Journal of Business & Economic Statistics* 37 (3): 447–456.

Gruber, J., and E. Saez. 2002. "The elasticity of taxable income: Evidence and implications." *Journal of Public Economics* 84 (1): 1–32.

Jones, D. 2012. "Inertia and overwithholding: Explaining the prevalence of income tax refunds." *American Economic Journal: Economic Policy* 4 (1): 158–185.

Kleven, H.J., and E.A. Schultz. 2014. "Estimating taxable income responses using Danish tax reforms." *American Economic Journal: Economic Policy* 6 (4): 271–301.

Kotlikoff, L.J., and D. Rapson. 2007. "Does it pay, at the margin, to work and save? Measuring effective marginal tax rates on Americans' labor supply and saving." *Tax Policy and the Economy* 21: 83–143.

Laurin, M., D. Messacar, and P.-C. Michaud. 2021. "Financial literacy and the timing of tax-preferred savings account withdrawals." Forthcoming in *Journal of Accounting and Public Policy (special issue on Accounting and Public Finance).*

Lavecchia, A. 2019. *'Back-loaded' tax subsidies for saving, asset location and crowd-out: Evidence from tax-free savings accounts.* Department of Economics Working Paper, no. 2019-04. Hamilton: McMaster University.

Liebman, J.B., and R.J. Zeckhauser. 2004 *Schmeduling*. Unpublished manuscript.

Mawani, A. and S. Paquette. 2011. "Pre-retirement RRSP withdrawals." *Canadian Tax Journal* 59 (2): 183–219.

Messacar, D. 2017. "Trends in RRSP contributions and pre-retirement withdrawals, 2000 to 2013." *Economic Insights*, no. 64. Statistics Canada Catalogue no. 11-626-X. Ottawa: Statistics Canada.

Messacar, D. 2018a. "Crowd-out, education, and employer contributions to workplace pensions: Evidence from Canadian tax records." *Review of Economics and Statistics* 100 (4) 648–663.

Messacar, D. 2018b. "The effect of tax withholding on pre-retirement savings withdrawals: Evidence from Canada." *Journal of Pension Economics and Finance* 17 (4): 534–553.

Milligan, K. 2009. *Time for a Tune Up: Personal Income Taxation in Western Canada.* Going for Gold Project Working Paper Series. Calgary: Canada West Foundation.

Milligan, K. 2016. Canadian Tax and Credit Simulator (CTaCS).

OECD 2012. *OECD Pensions Outlook 2012.* Paris: Organisation for Economic Co-operation and Development.

OECD 2013. *Pension Markets in Focus 2013.* Paris: Organisation for Economic Co-operation and Development.

Olsen, T. 2008. *BUNCH_COUNT: Stata module for plotting and quantifying bunching at kink points.*

Rees-Jones, A. 2018. "Quantifying loss-averse tax manipulation." *Review of Economic Studies* 85 (2): 1251–1278.

Rees-Jones, A., and D. Taubinsky. 2019. "Measuring 'schmeduling'." *Review of Economic Studies*. DOI: https://doi.org/10.1093/restud/rdz045.

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

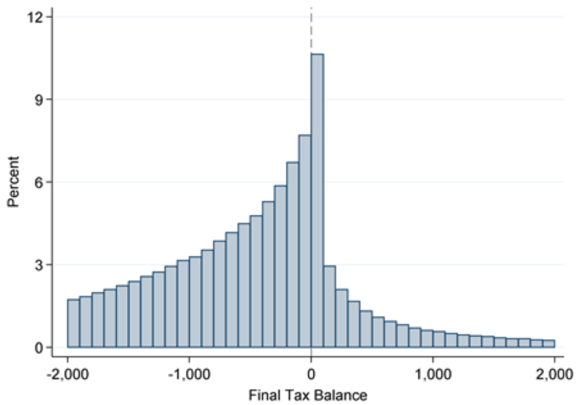
Søgaard, J.E. 2019. "Labor supply and optimization frictions: Evidence from the Danish student labor market." *Journal of Public Economics* 173: 125–138.

Stepner, M. 2013. "BINSCATTER: Stata module to generate binned scatterplots." *Statistical Software Components*, no. S457709. Boston: Boston College Department of Economics.

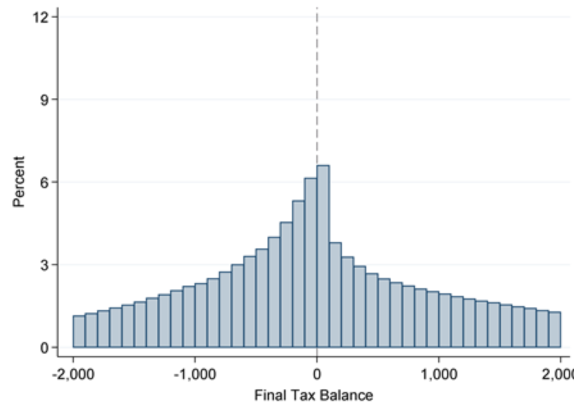
Stiglitz, J.E. 1985. "The general theory of tax avoidance." *National Tax Journal* 38 (3) 325–337.

Veall, M. 2001. "Did the tax flattening affect RRSP contributions?" *Canadian Journal of Economics* 34 (4): 120–131.

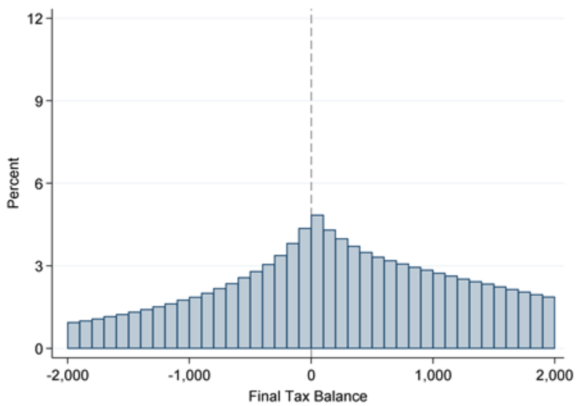
**Figure 1**
**Distributions of Actual and Adjusted Final Tax Balances—Tax Dataset**




(a) Actual Final Tax Balances in Full Sample



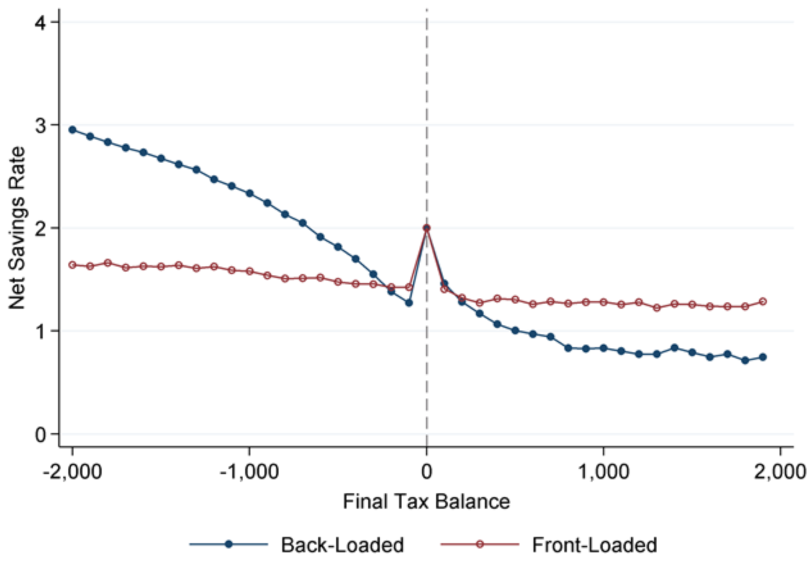

(b) Adjusted Final Tax Balances in Full Sample



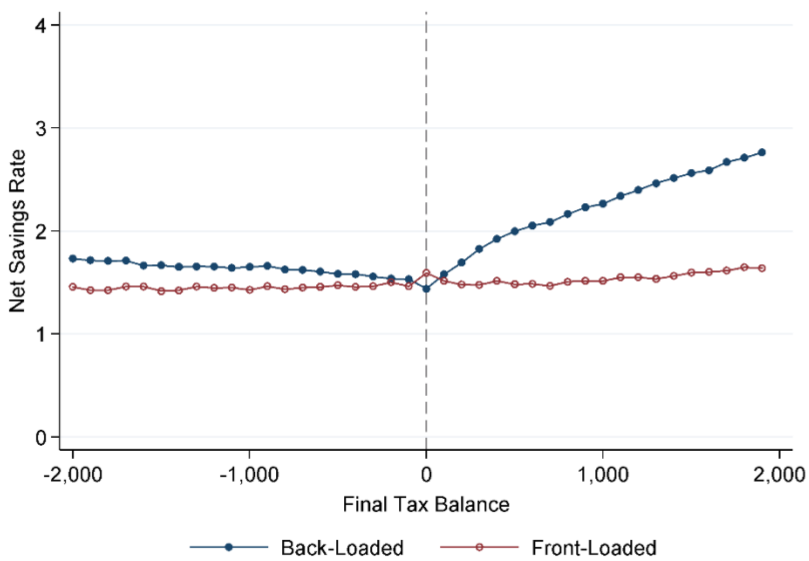

(c) Adjusted Final Tax Balances in Restricted Sample

**Notes:** Restricted to tax filers aged 20 to 54 with positive taxable incomes and tax withholdings, and the restricted sample also excludes those with no deductions; see the main text for further discussion. The adjusted final tax balance is the predicted amount owed to the central tax authority at the end of the year before deductions. Balances are expressed in 2016 constant dollars. Bins are of width $100.
**Source:** Statistics Canada, Longitudinal Administrative Databank, 2009 to 2016.

**Figure 2**
**Savings Rates and Final Tax Balances—Tax Dataset**



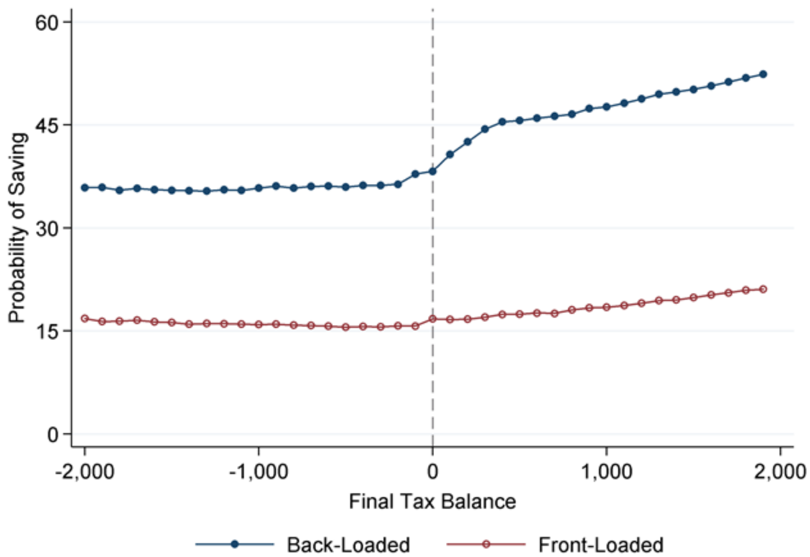

(a) Actual Final Tax Balances



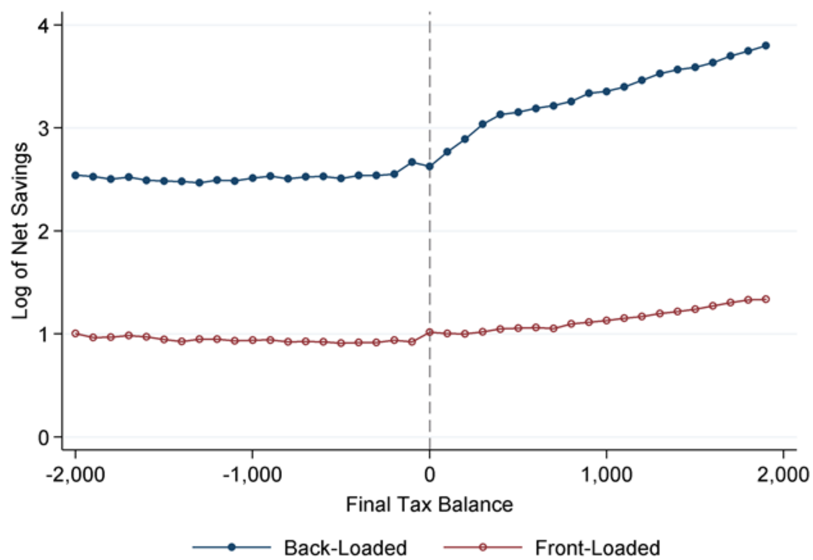

(b) Adjusted Final Tax Balances

**Notes:** Restricted to tax filers aged 20 to 54 with positive taxable incomes, tax withholdings and deductions. The adjusted final tax balance is the predicted amount owed to the central tax authority at the end of the year before deductions. The savings rates in back-loaded and front-loaded plans are expressed relative to total income before tax and after tax, respectively, since contributions to back-loaded plans are made with pre-tax income and contributions to front-loaded plans are made with after-tax income. Bins are of width $100.
**Source:** Statistics Canada, Longitudinal Administrative Databank, 2009 to 2016.

**Figure 3**
**Probability of Saving, Savings and Adjusted Final Tax Balances—Tax Dataset**



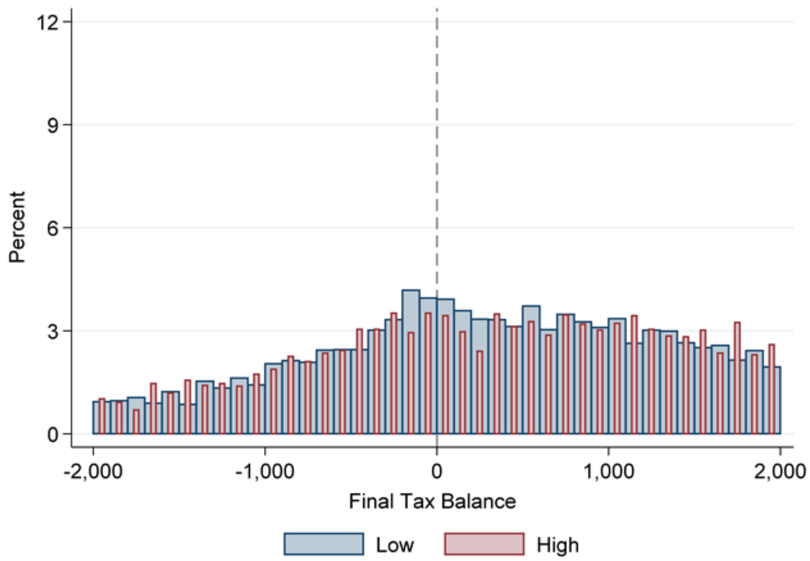

(a) Probability of Saving




(b) Log of Savings

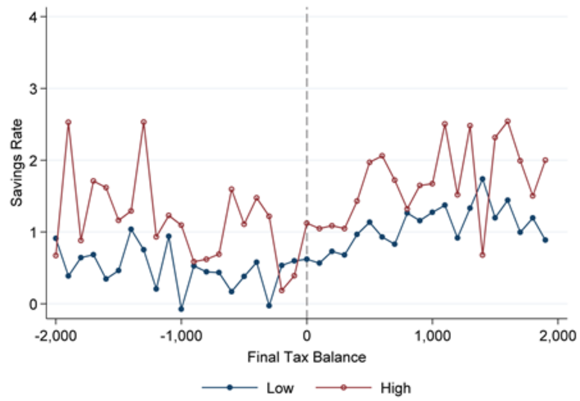
**Notes:** Restricted to tax filers aged 20 to 54 with positive taxable incomes, tax withholdings and deductions. The adjusted final tax balance is the predicted amount owed to the central tax authority at the end of the year assuming zero contributions to back-loaded savings plans (i.e., no tax deduction). Panel A shows the probability (from 0 to 100 percent) of having positive contributions each year. Panel B shows the log of savings. Bins are of width $100.
**Source:** Statistics Canada, Longitudinal Administrative Databank, 2009 to 2016.

**Figure 4**
**Distribution of Adjusted Final Tax Balances by Financial Literacy—Linked Dataset**

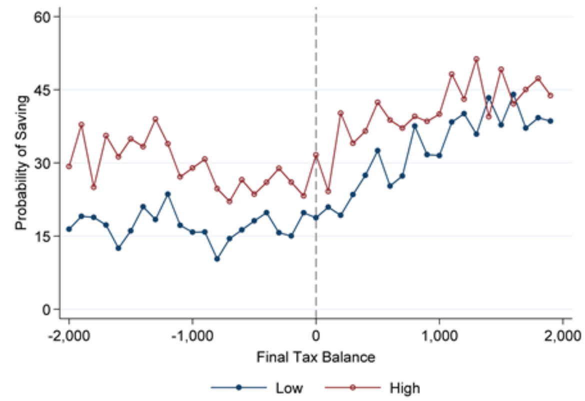



**Notes:** Restricted to tax filers aged 20 to 54 with positive taxable incomes, tax withholdings and deductions. The adjusted final tax balance is the predicted amount owed to the central tax authority at the end of the year before deductions. See the notes in Figure 1 and the discussion in the main text for more information. Balances are expressed in 2016 constant dollars. Bins are of width $100.
**Source:** Statistics Canada, Financial Capability, Employment and Income Database, 2009 to 2016.

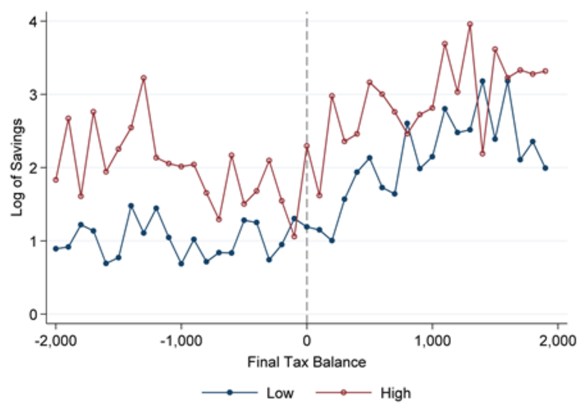
**Figure 5**
**Savings and Adjusted Final Tax Balances by Financial Literacy—Linked Dataset**




(a) Savings Rate




(b) Probability of Saving

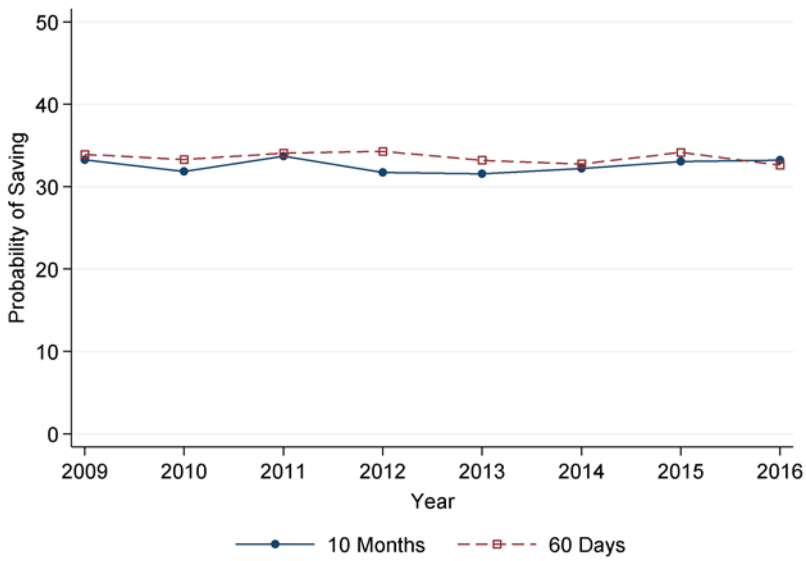



(c) Log of Savings

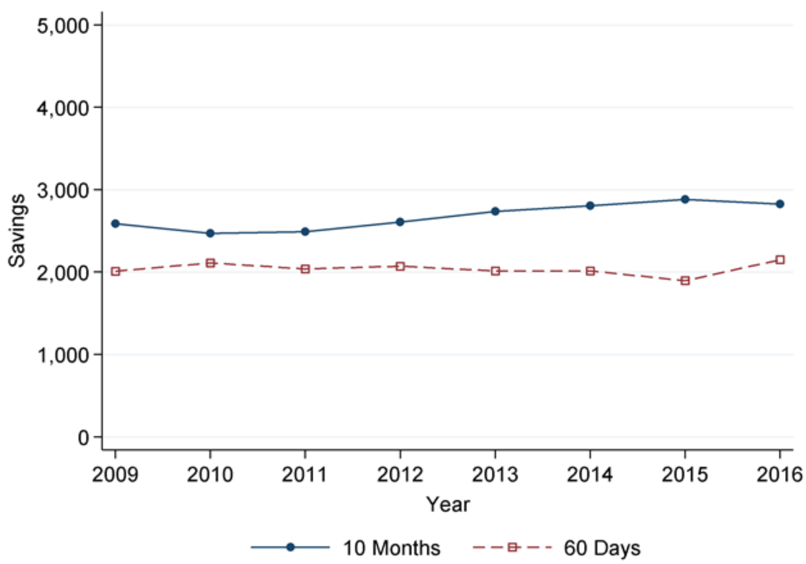
**Notes:** Restricted to tax filers aged 20 to 54 with positive taxable incomes, tax withholdings and deductions. The adjusted final tax balance is the predicted amount owed to the central tax authority at the end of the year before deductions. All currency values are expressed in 2016 constant dollars. See the notes in Figure 1 and the discussion in the main text for more information. Bins are of width $100.

**Source:** Statistics Canada, Financial Capability, Employment and Income Database, 2009 to 2016.

**Figure 6**
**Saving in Back-Loaded Plans in the 10 Months and 60 Days—Linked Dataset**




(a) Probability of Saving




(b) Conditional Savings

**Notes:** Restricted to tax filers aged 20 to 54 with positive taxable incomes, tax withholdings and deductions. Conditonal savings refers to the average value conditional on being strictly positive. All currency values are expressed in 2016 constant dollars.
**Source:** Statistics Canada, Financial Capability, Employment and Income Database, 2009 to 2016.

**Table 1**
**Descriptive Statistics—Tax and Linked Datasets**

| | Tax Dataset | | Linked Dataset | |
|---|---|---|---|---|
| | Full Sample (1) | Restricted Sample (2) | Full Sample (3) | Restricted Sample (4) |
| Demographics | | | | |
| Age (Years) | 37.4 | 38.9 | 39.4 | 39.3 |
| Female (Percent) | 49.2 | 49.7 | 50.8 | 50.8 |
| Married (Percent) | 56.5 | 62.6 | 57.7 | 57.4 |
| Educational Attainment (Percent) | | | | |
| High School Diploma or Less | | | 23.7 | 24.2 |
| Trades Certificate | | | 13.4 | 13.4 |
| Some Postsecondary or University Certificate | | | 33.9 | 34.0 |
| University Degree (Bachelor's or Higher) | | | 29.0 | 28.5 |
| Sources of Income (Percent) | | | | |
| Has Employment Income | 98.8 | 99.6 | 97.4 | 97.3 |
| Has Self-Employment Income | 5.1 | 6.8 | 5.7 | 5.6 |
| Has Investment/Dividend Income or Capital Gains | 16.1 | 19.5 | 17.5 | 17.3 |
| Has Unemployment Insurance Income | 16.8 | 16.6 | 18.0 | 18.1 |
| Conditional Income (2016 Constant Dollars) | | | | |
| Employment Income | 49,750 | 58,700 | 52,600 | 52,050 |
| Self-Employment Income | 13,550 | 14,100 | 14,350 | 14,350 |
| Investment/Dividend Income or Capital Gains | 7,850 | 7,350 | 7,100 | 7,150 |
| Unemployment Insurance Income | 7,550 | 7,600 | 7,950 | 8,000 |
| Sources of Saving (Percent) | | | | |
| Back-Loaded Plans | 32.0 | 46.2 | 37.4 | 36.4 |
| Front-Loaded Plans | 16.4 | 19.2 | | |
| Conditional Savings (2016 Constant Dollars) | | | | |
| Back-Loaded Plans | 4,750 | 4,750 | 4,050 | 4,050 |
| Front-Loaded Plans | 4,850 | 5,150 | | |
| Tax Liabilities | | | | |
| Federal Tax Liabilities (2016 Constant Dollars) | 5,550 | 6,900 | 5,900 | 5,800 |
| Provincial Tax Liabilities (2016 Constant Dollars) | 3,550 | 4,400 | 3,000 | 2,950 |
| Marginal Effective Tax Rate (Percent) | 30.4 | 33.2 | 32.2 | 32.0 |
| Number of Observations | 18,455,334 | 12,794,307 | 18,603 | 18,090 |

**Notes:** Restricted to tax filers aged 20 to 54 with positive taxable incomes, tax withholdings and deductions. The indicators for having various sources of income and saving are based on whether the value equals $100 or more. Conditional income refers to the level of income conditional on a strictly positive value. All currency values are expressed in 2016 constant dollars rounded to the nearest $50. The tax dataset does not contain educational attainment so this information is reported only for the linked dataset. The linked dataset does not contain savings information in front-loaded plans. The marginal effective tax rate is calculated using the Canadian Tax and Credit Simulator (CTaCS) created by Milligan (2016). The mean values are based on the full sample of observations; for example, the mean age refers to the average for all tax filers across all years observed.
**Source:** Statistics Canada, Longitudinal Administrative Databank ("Tax Dataset") and Financial Capability, Employment and Income Database ("Linked Dataset"), 2009 to 2016.

**Table 2**
**Estimated Relationship between Savings in Back-Loaded Plans and Adjusted Final Tax Balances—Tax Dataset**

| | Linear Trend, No Controls (1) | Quadratic Trend, No Controls (2) | Quadratic Trend, With Controls (3) | Quadratic Trend, With Individual Fixed Effects and Controls (4) |
|---|---|---|---|---|
| | Panel A: Savings Rate | | | |
| Change in Slope | 0.0771*** | 0.0635*** | 0.1240*** | 0.0659*** |
| | (0.0039) | (0.0137) | (0.0198) | (0.0115) |
| R-Squared | 0.0070 | 0.0344 | 0.0813 | 0.6017 |
| Observations | 8,338,849 | 8,338,849 | 8,338,849 | 7,693,801 |
| | Panel B: Probability of Saving | | | |
| Change in Slope | 0.5137*** | -0.2351 | 0.3687* | 0.3434*** |
| | (0.0888) | (0.1590) | (0.2046) | (0.1213) |
| R-Squared | 0.0135 | 0.0956 | 0.1483 | 0.7352 |
| Observations | 8,338,849 | 8,338,849 | 8,338,849 | 7,693,801 |
| | Panel C: Log of Savings | | | |
| Change in Slope | 0.0484*** | -0.0334** | 0.0255 | 0.0175 |
| | (0.0047) | (0.0160) | (0.0212) | (0.0122) |
| R-Squared | 0.0073 | 0.0705 | 0.1128 | 0.6642 |
| Observations | 8,338,849 | 8,338,849 | 8,338,849 | 7,693,801 |

**Notes:** Restricted to tax filers aged 20 to 54 with positive taxable incomes, tax withholdings and deductions. The control variables are fixed effects for year, age, gender, province, marital status, and employer-sponsored pension plan contributions; if individual fixed effects are included, then controls for age, gender and province are omitted. Standard errors are clustered by the forcing variable (adjusted final tax balances) in bins of width $100. ***, ** and * denote statistical significances at the 1, 5 and 10 percent levels, respectively.
**Source:** Statistics Canada, Longitudinal Administrative Databank, 2009 to 2016.

**Table 3**
**Estimated Relationship between Savings in Front-Loaded Plans and Adjusted Final Tax Balances—Tax Dataset**

| | Linear Trend, No Controls (1) | Quadratic Trend, No Controls (2) | Quadratic Trend, With Controls (3) | Quadratic Trend, With Individual Fixed Effects and Controls (4) |
|---|---|---|---|---|
| | Panel A: Savings Rate | | | |
| Change in Slope | 0.0027 | -0.0069 | 0.0080 | -0.0067 |
| | (0.0029) | (0.0140) | (0.0101) | (0.0041) |
| R-Squared | 0.0003 | 0.0011 | 0.0078 | 0.3664 |
| Observations | 8,338,849 | 8,338,849 | 8,338,849 | 7,693,801 |
| | Panel B: Probability of Saving | | | |
| Change in Slope | 0.2650*** | -0.1630** | -0.0260 | -0.0277 |
| | (0.0214) | (0.0727) | (0.0349) | (0.0217) |
| R-Squared | 0.0047 | 0.0133 | 0.0297 | 0.5188 |
| Observations | 8,338,849 | 8,338,849 | 8,338,849 | 7,693,801 |
| | Panel C: Log of Savings | | | |
| Change in Slope | 0.0201*** | -0.0176** | -0.0068 | -0.0079** |
| | (0.0019) | (0.0070) | (0.0044) | (0.0032) |
| R-Squared | 0.0015 | 0.0066 | 0.0147 | 0.3565 |
| Observations | 8,338,849 | 8,338,849 | 8,338,849 | 7,693,801 |

**Notes:** Restricted to tax filers aged 20 to 54 with positive taxable incomes, tax withholdings and deductions. The control variables are fixed effects for year, age, gender, province, marital status, and employer-sponsored pension plan contributions; if individual fixed effects are included, then controls for age, gender and province are omitted. Standard errors are clustered by the forcing variable (adjusted final tax balances) in bins of width $100. ***, ** and * denote statistical significances at the 1, 5 and 10 percent levels, respectively.
**Source:** Statistics Canada, Longitudinal Administrative Databank, 2009 to 2016.

**Table 4**
**Heterogeneous Effects of Financial Literacy on the Estimated Relationship between Savings in Back-Loaded Plans and Adjusted Final Tax Balances—Linked Dataset**

| | Unconditional | | Low Financial Literacy | | High Financial Literacy | |
|---|---|---|---|---|---|---|
| | No Individual Fixed Effects | With Individual Fixed Effects | No Individual Fixed Effects | With Individual Fixed Effects | No Individual Fixed Effects | With Individual Fixed Effects |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | Panel A: Savings Rate | | | | | |
| Change in Slope | 0.1195*** | 0.0806** | 0.0876 | 0.0486 | 0.1607* | 0.1435 |
| | (0.0436) | (0.0392) | (0.0522) | (0.0508) | (0.0895) | (0.0979) |
| R-Squared | 0.0546 | 0.5790 | 0.0426 | 0.5591 | 0.0752 | 0.5947 |
| Observations | 10,571 | 10,125 | 6,530 | 6,263 | 4,041 | 3,862 |
| | Panel B: Probability of Saving | | | | | |
| Change in Slope | 1.8526*** | 1.0396*** | 1.9309*** | 1.0979*** | 1.5142* | 0.8840 |
| | (0.5345) | (0.3081) | (0.6444) | (0.3970) | (0.8572) | (0.7291) |
| R-Squared | 0.1481 | 0.7419 | 0.1324 | 0.7269 | 0.1658 | 0.7573 |
| Observations | 10,571 | 10,125 | 6,530 | 6,263 | 4,041 | 3,862 |
| | Panel C: Log of Savings | | | | | |
| Change in Slope | 0.1791*** | 0.0873** | 0.1850** | 0.0547 | 0.1505 | 0.1381 |
| | (0.0535) | (0.0371) | (0.0714) | (0.0427) | (0.1002) | (0.0906) |
| R-Squared | 0.1065 | 0.6556 | 0.0826 | 0.6250 | 0.1395 | 0.6902 |
| Observations | 10,571 | 10,125 | 6,530 | 6,263 | 4,041 | 3,862 |

**Notes:** Restricted to tax filers aged 20 to 54 with positive taxable incomes, tax withholdings and deductions. The control variables are included in every regression. These controls include a quadratic trend in the forcing variable and fixed effects for year, age, gender, province, marital status, and employer-sponsored pension plan contributions; if individual fixed effects are included, then controls for age, gender and province are omitted. Standard errors are clustered by the forcing variable (adjusted final tax balances) in bins of width $100. ***, ** and * denote statistical significances at the 1, 5 and 10 percent levels, respectively.
**Source:** Statistics Canada, Financial Capability, Employment and Income Database, 2009 to 2016.

**Table 5**

**Heterogeneous Effects of Financial Literacy on the Estimated Relationship between Savings in Back-Loaded Plans and Adjusted Final Tax Balances Based on the Timing of Contributing Relative to the Deadline—Linked Dataset**

| | 10 Months | | | | 60 Days | | | |
|---|---|---|---|---|---|---|---|---|
| | Low Fiancial Literacy | | High Fiancial Literacy | | Low Fiancial Literacy | | High Fiancial Literacy | |
| | No Individual Fixed Effects | With Individual Fixed Effects | No Individual Fixed Effects | With Individual Fixed Effects | No Individual Fixed Effects | With Individual Fixed Effects | No Individual Fixed Effects | With Individual Fixed Effects |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | Panel A: Savings Rate | | | | | | | |
| Change in Slope | 0.0732** | 0.0774*** | 0.0512 | -0.0109 | 0.2804*** | 0.2502*** | 0.1290 | 0.1704 |
| | (0.0340) | (0.0244) | (0.0569) | (0.0350) | (0.0903) | (0.0837) | (0.1232) | (0.1105) |
| R-Squared | 0.0760 | 0.7435 | 0.0983 | 0.7361 | 0.0626 | 0.5124 | 0.0921 | 0.5259 |
| Observations | 6,530 | 6,263 | 4,041 | 3,862 | 6,836 | 6,596 | 4,353 | 4,165 |
| | Panel B: Probability of Saving | | | | | | | |
| Change in Slope | 0.9996 | 0.2098 | 0.6418 | -0.1831 | 2.0562*** | 0.4848* | 1.5144* | 0.2948 |
| | (0.6075) | (0.3549) | (0.9105) | (0.4099) | (0.6277) | (0.2583) | (0.7576) | (0.3500) |
| R-Squared | 0.1217 | 0.7764 | 0.1395 | 0.7797 | 0.2056 | 0.7636 | 0.2306 | 0.7564 |
| Observations | 6,530 | 6,263 | 4,041 | 3,862 | 6,836 | 6,596 | 4,353 | 4,165 |
| | Panel C: Log of Savings | | | | | | | |
| Change in Slope | 0.0815 | 0.0243 | 0.0705 | -0.0077 | 0.1496*** | 0.0411** | 0.1019* | 0.0261 |
| | (0.0490) | (0.0286) | (0.0737) | (0.0319) | (0.0445) | (0.0198) | (0.0520) | (0.0243) |
| R-Squared | 0.1377 | 0.7898 | 0.1573 | 0.7893 | 0.2269 | 0.7610 | 0.2594 | 0.7497 |
| Observations | 6,530 | 6,263 | 4,041 | 3,862 | 6,836 | 6,596 | 4,353 | 4,165 |

**Notes:** Restricted to tax filers aged 20 to 54 with positive taxable incomes, tax withholdings and deductions. The control variables are included in every regression. These controls include a quadratic trend in the forcing variable and fixed effects for year, age, gender, province, marital status, and employer-sponsored pension plan contributions; if individual fixed effects are included, then controls for age, gender and province are omitted. Standard errors are clustered by the forcing variable (adjusted final tax balances) in bins of width $100. ***, ** and * denote statistical significances at the 1, 5 and 10 percent levels, respectively.

**Source:** Statistics Canada, Financial Capability, Employment and Income Database, 2009 to 2016.

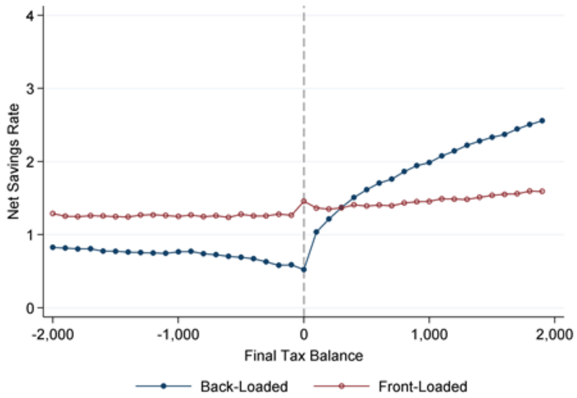
**Table 6**
**Effect of the Timing of Contributing on Next-Period Withdrawals—Linked Dataset**

| | Unconditional | | Low Financial Literacy | | High Financial Literacy | |
|---|---|---|---|---|---|---|
| | No Individual Fixed Effects | With Individual Fixed Effects | No Individual Fixed Effects | With Individual Fixed Effects | No Individual Fixed Effects | With Individual Fixed Effects |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | Panel A: Probability of Withdrawing | | | | | |
| 10 Months | 6.5457*** | 2.6331** | 9.1663*** | 3.6254** | 4.1156*** | 1.8769 |
| | (1.0138) | (1.1113) | (1.5471) | (1.5695) | (1.3096) | (1.5539) |
| 60 Days | -1.2200 | 0.0859 | -2.5122* | -1.8579 | -0.1429 | 1.9944 |
| | (0.9787) | (1.1512) | (1.4335) | (1.6099) | (1.3618) | (1.6394) |
| R-Squared | 0.0303 | 0.5405 | 0.0441 | 0.5264 | 0.0286 | 0.5618 |
| Observations | 15,355 | 15,114 | 8,678 | 8,537 | 6,677 | 6,577 |
| | Panel B: Log of Withdrawals | | | | | |
| 10 Months | 0.5495*** | 0.2298** | 0.7116*** | 0.2752** | 0.7116*** | 0.2752** |
| | (0.0835) | (0.0906) | (0.1252) | (0.1231) | (0.1252) | (0.1231) |
| 60 Days | -0.0749 | 0.0281 | -0.1542 | -0.0747 | -0.1542 | -0.0747 |
| | (0.0800) | (0.0922) | (0.1167) | (0.1235) | (0.1167) | (0.1235) |
| R-Squared | 0.0340 | 0.5172 | 0.0464 | 0.4985 | 0.0464 | 0.4985 |
| Observations | 15,355 | 15,114 | 8,678 | 8,537 | 8,678 | 8,537 |

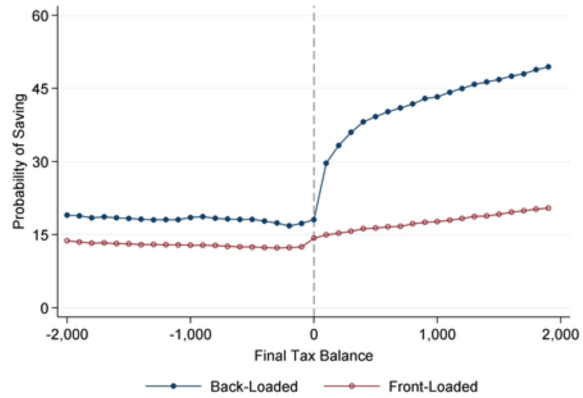
**Notes:** Restricted to tax filers aged 20 to 54 with positive taxable incomes, tax withholdings and deductions. The control variables are included in every regression. These controls include a quadratic trend in the forcing variable and fixed effects for year, age, gender, province, marital status, and employer-sponsored pension plan contributions; if individual fixed effects are included, then controls for age, gender and province are omitted. Standard errors are clustered by the forcing variable (adjusted final tax balances) in bins of width $100. ***, ** and * denote statistical significances at the 1, 5 and 10 percent levels, respectively.
**Source:** Statistics Canada, Financial Capability, Employment and Income Database, 2009 to 2016
**Notes: Source:** Statistics Canada, Longitudinal Administrative Databank, 2009 to 2016.

**Appendix**

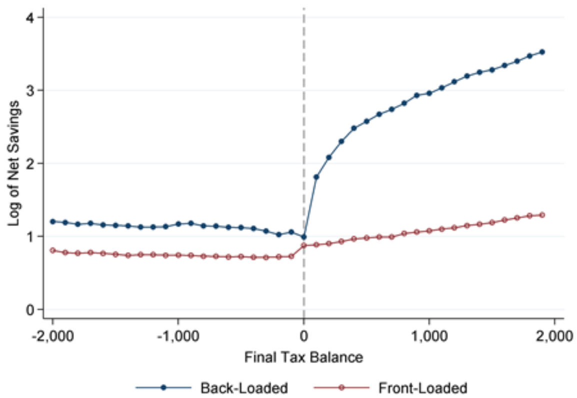
**Figure A1**
**Relationship between Saving and Adjusted Final Tax Balances in the Full Sample—Tax Dataset**




(a) Savings Rate




(b) Probability of Saving




(c) Log of Savings

**Notes:** This figure replicates the analysis in Panel B of Figure 2 and Figure 3 but using the full sample of tax filers including those without deductions. All currency values are expressed in 2016 constant dollars. See the notes in Figures 2 and 3 and discussion in the main text for more information.
**Source:** Statistics Canada, Longitudinal Administrative Databank, 2009 to 2016.

**Table A1**
**Actual versus Predicted Total Tax Liabilities—Tax Dataset**

| | Actual<br>(1) | Predicted<br>(2) |
|---|---|---|
| Full Sample | 8,100 | 10,000 |
| By Age Group | | |
| Less Than 40 | 6,500 | 8,300 |
| 40 or More | 10,250 | 12,200 |
| By Gender | | |
| Female | 6,350 | 8,000 |
| Male | 9,850 | 12,000 |
| By Marital Status | | |
| Married | 8,950 | 10,850 |
| Unmarried | 7,950 | 9,800 |
| By Employment Status | | |
| Employed | 8,200 | 10,100 |
| Self-Employed | 7,050 | 9,000 |
| Neither Employed nor Self-Employed | 1,450 | 1,550 |
| By Savings | | |
| Has Savings in Back-Loaded Plans | 12,500 | 14,900 |
| Has Savings in Front-Loaded Plans | 9,900 | 12,150 |
| No Tax-Preferred Savings | 5,950 | 7,550 |

**Notes:** Tax filers who are married include both legally married and common-law relationships. Employment and self-employment status are defined based on whether the tax filer has any employment and self-employment income for the year, respectively. To account for outliers since a predicted value is used, the statistics exclude observations above the $99^{th}$ percentile. All currency values are expressed in 2016 constant dollars.
**Source:** Statistics Canada, Longitudinal Administrative Databank, 2009 to 2016, and the Canadian Tax and Credit Simulator (CTaCS) created by Milligan (2016).