Canadian Labour Market and Skills Researcher Network

Working Paper No. 146

Taxation and top incomes in Canada

Kevin Milligan
University of British Columbia

Michael Smart
University of Toronto

November 2014

CLSRN is funded by the Social Sciences and Humanities Research Council of Canada (SSHRC) under its Strategic Knowledge Clusters Program. Research activities of CLSRN are carried out with support of Human Resources and Skills Development Canada (HRSDC). All opinions are those of the authors and do not reflect the views of HRSDC or the SSHRC.

Taxation and top incomes in Canada [*]

Kevin Milligan[†]
and
Michael Smart[‡]

August, 2014

Abstract

We estimate the elasticity of reported income with respect to tax rates for high earners using sub-national variation across Canadian provinces. We argue this allows for better identification of tax elasticities than the existing literature. We find that elasticities of reported income at the provincial level are large for incomes in the top one percent, but small for lower earners. There are strong indications that the response happens both through earned and capital income. While our estimated elasticities are large, changes in tax rates cannot explain much of the overall long-run trend of higher income concentration in Canada.

JEL Codes: H21, H24, D31
Keywords: income concentration, income taxation, taxable income elasticity

[*] We thank seminar participants at Carleton University, Exeter University and University of Ottawa for comments on this project.

[†] University of British Columbia, kevin.milligan@ubc.ca
[‡] University of Toronto, msmart@chass.utoronto.ca

1 Introduction

An enduring controversy in the study of high earners is the role of taxation in understanding the evolution and composition of income at the top of the income distribution. Piketty and Saez (2003) uncovered a strong trend in concentration of incomes in the highest fractiles in the United States starting in the 1980s. That decade also saw large cuts to marginal tax rates faced by higher earners. While these two trends lined up quite well in the United States, evidence from other countries has been less persuasive for the link between the surge in income concentration and high income taxation. Looking at the case of Canada, evidence from Saez and Veall (2005) and Veall (2012) suggests much less correspondence between taxes and the evolution of high incomes.

One of the challenges in this literature has been the necessary reliance on time series trends to make inferences. Given the national basis for most income tax systems, it has proven difficult to separate the impact of tax changes from other developments in the economy. Moreover, the usual route of exploiting large reforms within a country to make inferences is difficult in this case because the tax-response parameter being estimated may itself be different before and after the reform.[1]

In this paper we estimate the elasticity of reported income using the sub-national variation across Canadian provinces. We argue this allows for better identification of elasticities than the existing literature. There are two primary advantages. First, we have data and variation in tax rates within one national economic space, which allows us to control for common trends in a more rigorous way than in much of the literature. Second, we have an extended opportunity, both in time and across provinces, with a very similar tax base. This is important because theory and evidence suggests that the response of incomes to taxation should vary when the tax base changes. Anchoring our estimates in a period of tax base stability allows us to be more sure of the parameter we are estimating.[2]

An important limitation of our approach also results from our use of provincial-level variation. If some of the response to higher tax rates at the provincial level reflect movements of taxable dollars within-Canada across provinces, then the elasticity we estimate here based on provincial-level variation

[1] The Canadian tax reform of 1988 was used as the basis for estimation in Sillamaa and Veall (2001). However, Fullerton (1996) argues that times of large tax reforms are especially difficult for identifying the impact of taxation on one particular element of behaviour because so many things change at once.

[2] Our strategy aligns well with the suggestion in the survey by Saez, Slemrod and Giertz (2012) to seek "better sources of identification; for example, parallel income tax systems that differentially affect taxpayers over a long period of time."

cannot be applied without reservation to predict the impact of federal tax policy changes. The provincial estimates may embed some of this inter-provincial response that would not be present with a federal tax rate change. In related and ongoing research, Milligan and Smart (2013) develops a theoretical and empirical framework for evaluating the complex federal interactions in considering the federal and provincial aspects of high income taxation. The limitations of the approach in this present paper, though, mean that the estimates presented here are most useful for understanding changes at the provincial level.

Comparing across provinces and through time, we find that elasticities are large for incomes at the top of the income distribution, but small elsewhere. We also find that the overall long-run trend of higher income concentration in Canada is largely unrelated to taxes, as the long-run changes in tax rates are not large enough to explain the big changes in income concentration. Finally, we look at the composition of income, finding evidence that both earnings and capital income respond strongly to taxation.

We begin with an overview of the relevant institutions in Canada. This is followed by an extensive discussion of our empirical strategy. We then describe the data, before heading into our results. We conclude with a short discussion.

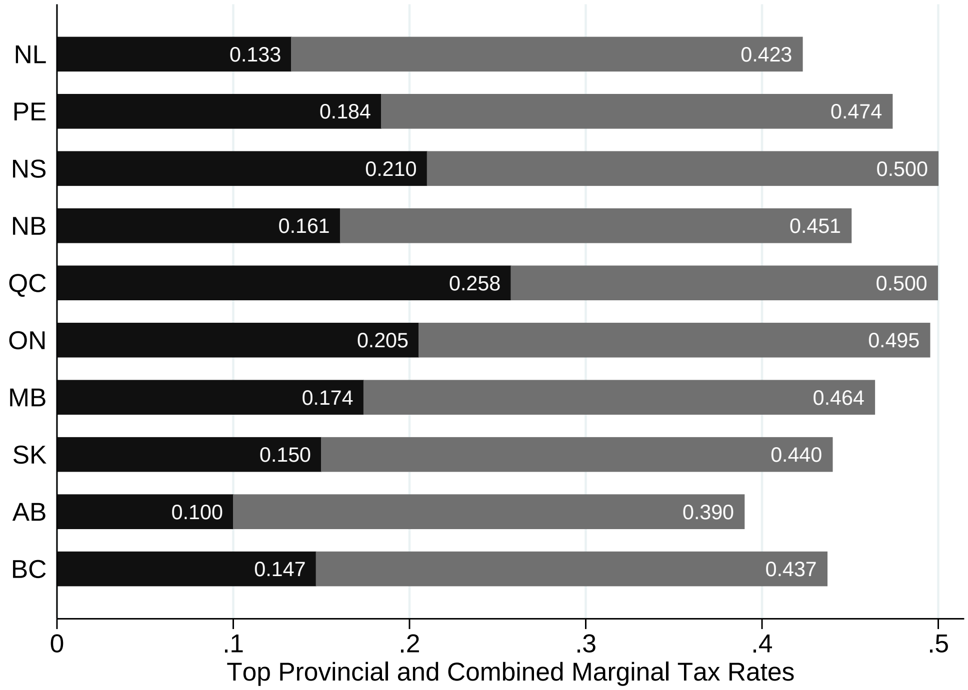
2 Personal income taxation in Canada

In the Canadian federation, the constitutional division of powers allows both the federal and provincial governments to assess direct taxation on personal incomes. Each province and the federal government set tax brackets and rates based on a definition of taxable income, with various credits applied against the calculated tax liability. For the provinces, this is called the 'tax on income' system, as provincial taxes (outside Quebec) are levied on federally-defined taxable income. In 2009, 39.6 percent of the total \$189.2 billion in income tax revenue in Canada was collected by provincial or territorial governments.[3] So, provincial income taxation in Canada is sizeable.

Nine of the provincial governments collect personal income taxes through tax collection agreements with the federal government, with only Quebec operating its own income tax system. Under these tax collection agreements, the federal government collects the income tax on behalf of the nine 'agreeing'

[3] Source is CANSIM series 385-0001. Data beyond 2009 are not available for personal income tax revenues in the new Government Finance Statistics series in table 385-0032.

Figure 1: Top Marginal Tax Rates, 2013



Source: Canadian Tax and Credit Simulator.

provincial governments. The tax collection agreements also require the use of the federal definition of taxable income. As noted in the introduction, this common tax base is useful both for estimation and for interpretation of the sensitivity of taxable income to tax rates.

The provincial tax rates for high earners vary strongly across the country, ranging from a low of 10 percent in Alberta to a high of 25.75 in Quebec. Figure 1 displays the provincial and combined marginal tax rates for 2013. The federal income tax rate for the highest bracket is 29 percent outside of Quebec, but drops to 24.2 percent in Quebec because of a federal tax abatement for residents of that province. The combined rate is lowest in Alberta (at 39 percent) and highest in Nova Scotia (at 50.0 percent).

This personal income tax system took its current 'tax on income' form in 2000/2001. Previous to then, the bulk of provincial income tax in the agreeing provinces was levied using a 'tax on tax' structure. Under 'tax on tax', the provincial liability was not calculated directly on federal taxable income, but as a percentage of a measure of federal taxes called Basic Federal Tax.[4] A province did not choose its own brackets and rates, just one flat percentage that was applied to Basic Federal Tax. In 1999, this rate varied from a high of 69 percent in Newfoundland and Labrador to a low of 39.5 percent in Ontario. So, if a taxpayer had $100 of Basic Federal Tax, the initial provincial tax liability (before surtaxes) was $39.50 in Ontario and $69.00 in Newfoundland and Labrador. This structure limited to some degree the scope of progressivity a province could apply to its tax schedule. However, provinces had the ability to add surtaxes and 'flat taxes' levied against taxable income.[5]

In Figure 2, we plot the provincial high income tax rate for each province for the years 1988 to 2011. We calculate this using the tax rate at the national income threshold of the 99.99th percentile income group, which is far above the top tax bracket threshold for each of the years and provinces included. It is difficult in places to pick out the different provincial lines, but this is a virtue for our identification strategy which relies on substantial variation across provinces. At the bottom of the graph, Alberta is a clear outlier after 1991 with rates substantially lower than the other provinces. British Columbia went from the lowest in 1988 to among the highest in the mid-1990s and then back to the second lowest in 2002. In all provinces, the 'tax on income' reform put in place in 2000-2001 is evident. In the 2000s, British Columbia moves from among the highest to the lowest outside Alberta, while Nova Scotia and Newfoundland and Labrador move in opposite directions.

We chose the year 1988 as the starting point for Figure 2 because 1988 was the year of a major reform to the tax base. Previous to 1988, the tax base was smaller with deductions from total income for many items that became credits in 1988.[6] This motivates our decision to begin the analysis with the year 1988. While there have been some base adjustments in the period since 1988, we view the tax base during this 1988 to 2011 era as relatively stable.[7]

One way to assess the extent of the province-year variation in the tax rates is to run a simple regres-

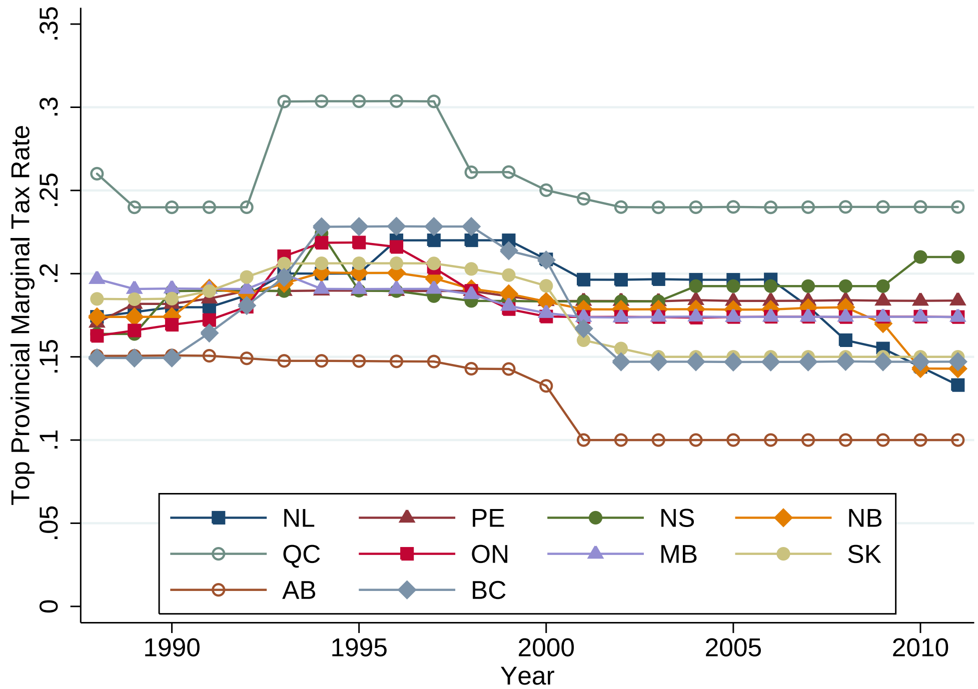
[4]Basic Federal Tax results from taxable income being applied to the federal bracket and rate structure, before calculating any surtaxes or abatements.

[5]That same Ontario taxpayer in 1999 was subject to surtaxes of 20 percent of provincial tax for each dollar of provincial tax over $3,750 and an additional 36 percent for each dollar of provincial tax over $4,681. Also in 1999, Manitoba, Saskatchewan, and Alberta had flat taxes ranging from 0.5 percent to 2.0 percent on taxable income.

[6]For example, the basic amount, the amount for Canada Pension Plan contributions, and the amount for medical expenses (among many others) were all transformed from deductions to credits.

[7]The most important changes to the tax base over this time period involve capital income, with the capital gains exclusion rate changing from 75 percent to 50 percent in 2000 and the dividend tax credit and gross-up also changing through time.

Figure 2: Top Provincial Marginal Tax Rates, 1988 to 2011



Source: Canadian Tax and Credit Simulator.

sion of the top marginal tax rate on year and province dummies and record the R-squared to see how much of the variation in tax rates is within province through time. Running this regression on the whole 1988 to 2011 dataset yields an R-squared of 80.9 percent, leaving a fairly substantial 19.1 percent of the variation explained by within-province movements through time. The R-squared is lower in the 'tax on tax' 1988-1999 period (at 71.6 percent) than in the 'tax on income' 2001-2011 period (at 88.2 percent). There is also more within-province variation in the five most easterly provinces (with an R-Squared of 65.8 percent) than in the five more westerly provinces (with an R-Squared of 88.2 percent).

3 Empirical Approach

We employ an aggregate shares approach to estimating the elasticity of reported income.[8] Very similar methods to our approach have been used in Saez (2004), Saez and Veall (2005), Atkinson and Leigh

[8] The different approaches taken in the literature are reviewed by Saez et al. (2012).

(2010), and Department of Finance (2010). In this section, we construct our empirical specification and discuss how our approach treats empirical challenges that arise.

The basic economic relationship we wish to estimate is between reported total income y_{it} of an individual i in year t and the marginal tax rate he or she faces τ_{it}. Since the work of Feldstein (1995), the basic empirical specification for estimating the elasticity of taxable income often takes the form

$$\log y_{it} = \alpha_i + \delta_t + \beta \log(1 - \tau_{it}) + \epsilon_{it} \tag{1}$$

where α_i and δ_t are individual and year fixed effects affecting mean reported income, β is the elasticity of reported income to the net-of-tax rate, and ϵ_{it} is an IID error term. This model has frequently been estimated with panel data on individual incomes using quasi-experimental methods, exploiting tax reforms that change the tax rates facing some taxpayers but not other taxpayers, who serve as a "control" group for estimating the unobserved time fixed effects δ_t driving income variation contemporaneous with the tax reforms. However, Saez et al. (2012) point out several challenges confronting panel data studies, including the availability of long panels encompassing repeated tax reforms, and the problem of disentangling mean reversion after temporary income shocks from changes to the income distribution caused by taxes.

An alternative approach is to estimate elasticities using only aggregate data on the income shares of fractiles of the income distribution over time. For simplicity, let us divide the population into a group of high income taxpayers H, all facing a common tax rate τ_{Ht}, and a complementary group of lower-income taxpayers L. The H group is some top fractile of the distribution, say the top one percent. For reasons that will become clear, we are primarily interested in estimating the elasticity of reported income β for the sub-population of H taxpayers, but the L group remains useful as a control for unobservable factors affecting incomes over time. In particular, to identify β using the share data approach, we assume that: (i) reported incomes in group H have a common tax elasticity β, but incomes in group L are not affected by tax rates; and (ii) the two groups are affected by temporal shocks in the same way on average—i.e. the usual parallel trends assumption.

Letting Y_{Ht} be aggregate income in the top fractile in year t and Y_{Lt} be aggregate income of others, the underlying model is therefore

$$\log Y_{Ht} = \alpha_H + \delta_t + \beta \log(1-\tau_{Ht}) + \epsilon_{Ht} \tag{2}$$

$$\log Y_{Lt} = \alpha_L + \delta_t + \qquad\qquad\qquad + \epsilon_{Lt}. \tag{3}$$

Formally, this specification corresponds to the individual income model (1) only if membership in the top fractile H does not vary from year to year, i.e. there are no income ranking reversals. However, Saez (2004) argues that if the goal is studying the tax-induced change to the income distribution rather than to individuals, then the elasticity estimated from share data is the relevant one.

Given time series data on aggregate incomes of the two groups and top tax rates, we can therefore difference the two equations to eliminate time effects and estimate β by OLS regression of $\log Y_{Ht} - \log Y_{Lt}$ on the tax rate term. As a practical matter, it is more useful to define the income share of the top fractile to be

$$s_{Ht} = \frac{Y_{Ht}}{Y_{Ht} + Y_{Lt}} \tag{4}$$

and the log share ratio

$$z_{Ht} \equiv \log\left(\frac{s_{Ht}}{1-s_{Ht}}\right) = \log s_{Ht} - \log(1-s_{Ht}) = \log Y_{Ht} - \log Y_{Lt} \tag{5}$$

From (5) we see that fixed effects estimation of our model (2)-(3) is equivalent to the simple OLS regression model

$$z_{Ht} = \alpha + \beta \log(1-\tau_{Ht}) + u_t \tag{6}$$

which is the basis for regressions reported in this paper.[9]

As noted, a key assumption in deriving (6) is that the control group is unaffected by taxes, or that the control group's own tax rate is unchanged in the regression sample. Previous research and our own investigations with the data suggest that behavioural response to taxes is much larger in the top one percent of the income distribution and above than among any lower fractiles. (See below for evidence

[9] In contrast, most of the literature surveyed in Saez et al. (2012) uses $\log s_{Ht}$ as the dependent variable rather than z_{Ht}. That regression consistently estimates β under slightly different and arguably less plausible identifying assumptions. It turns out that the differences in estimates under the two approaches is small for our data.

on elasticity heterogeneity; Saez et al. (2012) page 26 for a discussion of the assumption of zero response in the control group.) So, while the tax rate of the control group L might be included in (6), for efficiency reasons we choose to omit it entirely.[10] Furthermore, tax reforms in our sample period have resulted in larger variability of marginal tax rates for the top one per cent than for the lower fractiles. As one measure, the standard deviation of the marginal tax rate at the 50th percentile is 0.0294, but 0.0358 at the 99th percentile. Taking a taxfiler-weighted national average, the marginal tax rate at the 50th percentile ranged between 30 and 34 percent over 1988 to 2011, while at the 99th percentile the range was from 45 percent to 53 percent.

There remains the usual endogeneity concern that top tax rate changes may be correlated with unobservable factors driving income distribution over time so that $\text{corr}(\tau_{Ht}, u_t) \neq 0$ and OLS on (6) is biased. While there is much debate in the literature about the contribution of taxes to increasing income concentration, it seems extreme to assume all movements in income concentration are tax-driven. Solutions to this challenge in the literature have included trying to specify and include the non-tax factors that may drive trends in the top fractile share, or to include general polynomial time trends. For example, Saez (2004) uses time trends, while Saez and Veall (2007) uses the top one percent income share from the United States.

One approach to this problem that is possible with Canadian data is to exploit tax differences among Canadian provinces. If the non-tax determinants of high income concentration that are correlated with tax rate trends are common to all provinces, then the cross-province difference-in-difference estimator purges the national trends and looks instead at changes in tax rates within provinces over time. Thus we introduce a further set of time fixed effects common to all provinces into the income share equation equation (6), as well as a set of jurisdiction dummies to estimate

$$z_{Hpt} = \alpha_p + \lambda_t + \beta \log(1 - \tau_{Hpt}) + u_{pt}. \tag{7}$$

This estimator was also adopted in Atkinson and Leigh (2010), who estimate elasticities using time-series data on national top fractile shares in the English-speaking developed countries.[11] Arguably, our

[10] We have also run specifications that explicitly control for the marginal tax rate at the median earnings level and found the top elasticity changes little. We report the relevant coefficients in the discussion of Table 3 below.

[11] That is, the US, UK, Canada, Australia , and New Zealand.

approach using provincial data is preferable for two reasons. First, many of the time-varying institutional factors that lead to a concern about the correlation of u_t and τ_{Ht} are not common across countries. For example, wage-setting patterns or executive compensation practices have strong country-specific features. Second, it is not clear β is the same across countries. As noted by Slemrod and Kopczuk (2002), the elasticity of response to taxes will depend on the particular tax base. If a country's tax system allows easy shifting of income out of the personal income tax base, then the elasticity will be high; if the tax base is tighter the elasticity will be low. The elasticity is not a structural behavioural parameter, but a reflection of a particular tax system.

By using Canadian provinces to estimate equation (7), we get the benefits of the Atkinson-Leigh approach, but can improve on both of the shortcomings noted above. First, we argue that the non-tax determinants of high income concentration listed above are more likely to be common across provinces within Canada than they are across countries. Second, the tax bases across nine of the ten provinces are identical through the tax collection agreements, meaning we are estimating the response in the context of the same tax base.

The same province-year estimation strategy was also used in Department of Finance (2010). They implement the same basic strategy over the 1995-2006 period. In some specifications, they augment their analogue to equation (7) with a 'high oil price' dummy for certain years for Alberta. In some robustness checks, we include province-specific time trend polynomials that control (up to a polynomial) for such province-specific episodes in any province.

There are some limitations to this approach however. Our estimated elasticity parameter is most relevant from the point of view of a provincial government. If income shifting across provinces is empirically important, then a unilateral increase in the tax rate in one province will cause a larger decline in the tax base there than in the nation as a whole. Thus the elasticity of the provincial tax base may exceed the elasticity of the federal base. In the presence of interprovincial income shifting, a tax increase in one province may therefore have offsetting effects on reported income in the home province and in the other provinces that implicitly serve as the control group for our difference-in-difference estimator, which would tend to bias our estimated elasticity upward.

Our estimates would also be affected if individual taxpayers are physically mobile between provinces or across countries in response to tax rate changes. Our income share approach relies on changes in the income share of top fractiles within each province over time, but does not account for changes in the identity of those taxpayers. If, for example, some high income taxpayers were to leave a province in response to a tax rate increase, the right reported income to record for those who shifted residence would be zero, as those taxpayers responded to the tax change by moving all of their reported income out of that jurisdiction. However, in our approach those who depart are "replaced" in the data by taxpayers with incomes right at the fractile threshold. This understates the true decline in provincial income in response to the tax change. Therefore, with interjurisdictional mobility, we tend to under-estimate the tax base elasticity under our approach. Thus the two biases are in offsetting directions.

4 Data

We combine various sources of information to form our dataset. Below we describe the data sources in detail, then provide some graphical analysis of the broad trends in our data.

4.1 Sources

Our empirical strategy relates the the tax rate faced by top earners to the income reported by the top earners. So, our analysis requires both a measure of tax rates and measures of incomes. Below we describe how we use the Canadian Tax and Credit Simulator (CTaCS; see Milligan 2013a) to form our tax rates. We then describe our manipulation of the CANSIM high income database to form our measures of top incomes.

The CTaCS calculator produces a tax liability for an individual given a certain income, year, and province of residence using the tax system parameters in effect during that year in the given province. By repeating the calculation with an increment of $100 to the income, a marginal tax rate can be estimated by taking the difference of the two tax liabilities and dividing by 100. We perform this calculation for each year between 1988 and 2011 and for each province. For incomes, we perform the calculations both for the average and for the threshold level of each income group under consideration. For example, for the 'top 1 percent' income group, we calculate the marginal tax rate at both the income threshold for entry into the income group, and the average income among those in the income group.

Our focus on top earners means that Canada's vast and important system of refundable tax credits does not play a large role in our analysis. For middle and lower income Canadians, proper attribution of tax liabilities and calculation of marginal tax rates requires careful consideration of refundable tax credits like the Goods and Services Tax Credit and the Canada Child Tax Benefit. However, at the income ranges necessary for membership in the top fractiles of interest here, individuals are not eligible for the refundable tax credits.

The simulation of individual taxation at higher income ranges does require more attention be paid to deduction and credit items that may reduce taxable income or tax paid. We impute the most common of these items to our individuals using available data.[12] However, these imputations will only affect our estimates of marginal tax rates to the extent that the simulated individual is pushed into a lower tax bracket. Given the relatively low thresholds to be in the highest tax bracket in Canada (all members of the top one percent are in the highest tax bracket in most province-years), being pushed into a lower tax bracket is uncommon. Overall tax liabilities may be affected by these imputations, but marginal tax rates are unlikely to be affected by the imputations at higher fractiles.

Our income data come from the CANSIM high income database, which reports various measures of income for several income fractiles.[13] The CANSIM data distinguish between 'market income' and 'total income', which differ by the inclusion of government transfer income in total income, but not in market income. The other main distinction is whether or not to include capital gains income. We focus our analysis for the most part on total income excluding capital gains, but show sensitivity to the other income definitions as well. We transform all income data to 2010 values using the Consumer Price Index. The data are available for various income groups, including for those above cutoffs at the 90th, 95th, 99th. and 99.9th percentiles of income.

For our last set of results which look at the composition of income among high earners, we assemble a dataset based on the annual Canada Revenue Agency Income Statistics publications. These data report

[12] We impute based on data from the Canada Revenue Agency's Tax Statistics on Individuals. We define cells by province, year, and narrow income groups. We impute an amount and a probability of any amount to each cell based on these Canada Revenue Agency data. We include the following tax measures: donations and gifts, RRSP contributions, RPP contributions, union dues, childcare expenses, other deductions, and additional deductions from net income. We repeat this simulation 100 times and average the results.

[13] The relevant CANSIM table is 204-0002. This table provides average, median, and threshold cutoff income levels (among other statistics) for different income groups with a particular focus on high income groups. These data come from the Longitudinal Administrative Databank which is constructed from income tax administrative data.

the aggregate amount and number of taxfilers for a large set of items appearing on the tax form, ranging from income items like Total Income, taxable dividend income, and self-employment income to deduction items like childcare expenses and credit items like charitable donations. The data are reported in aggregate, as well as for a set of nominal categories which is fairly constant through time. We pay particular attention to the categories for $250,000 and those that report income between $150,000 and $250,000 as that income range is the closest to the P99 cutoff on which we focus. The appendix describes more detail about the formation of this dataset.

4.2 Income Trends

The trends in high incomes in Canada have been well-documented using administrative tax data in Saez and Veall (2005) and Veall (2012), and in census data by Milligan (2013b). In this section we provide a brief overview of the trends with a particular emphasis on the provincial attributes and the time period 1988 to 2011 of interest to our analysis.

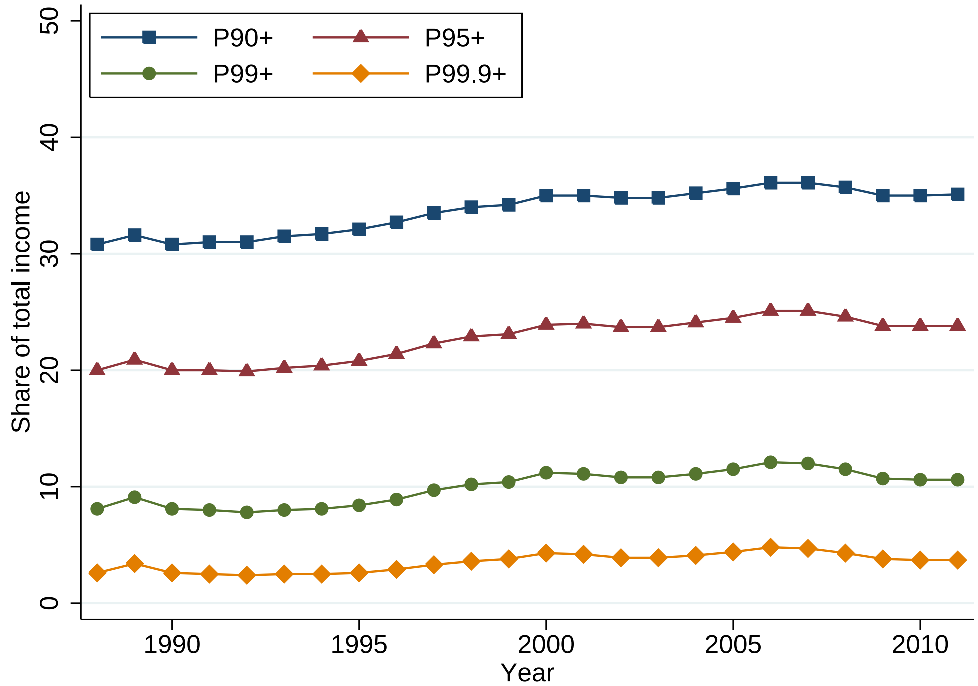
Table 1: Thresholds for high income groups, 1988 and 2011

	1988			2011		
	P95	P99	P99.9	P95	P99	P99.9
Canada	88,520	154,133	444,893	105,229	203,656	677,430
NL	71,667	108,810		98,039	168,094	
PE	70,031	114,046		81,715	139,722	
NS	78,212	129,099	333,792	90,266	154,588	399,831
NB	74,121	116,173		86,282	147,010	363,395
QC	81,648	133,190	316,612	91,140	169,649	505,254
ON	95,720	179,495	579,882	108,047	215,316	778,189
MB	78,376	125,336	310,394	91,917	161,098	469,595
SK	79,357	125,663	288,796	107,269	180,240	507,878
AB	92,284	160,515	416,422	138,945	281,096	1,044,419
BC	89,502	157,897	457,001	102,411	190,151	650,418

Source: CANSIM 204-0002. All dollar values converted to 2010 dollars using Consumer Price Index. Shown is the income cutoffs for the noted high income groups.

We begin with Table 1 which displays the cutoffs for the percentile income groups $\{P95, P99, P99.9\}$. We use total income without capital gains converted to 2010 dollar values here, and as the default for all our analysis. We show the cutoff values for 1988 and 2011 for Canada as a whole, and then separately by

Figure 3: Top income shares, 1988-2011



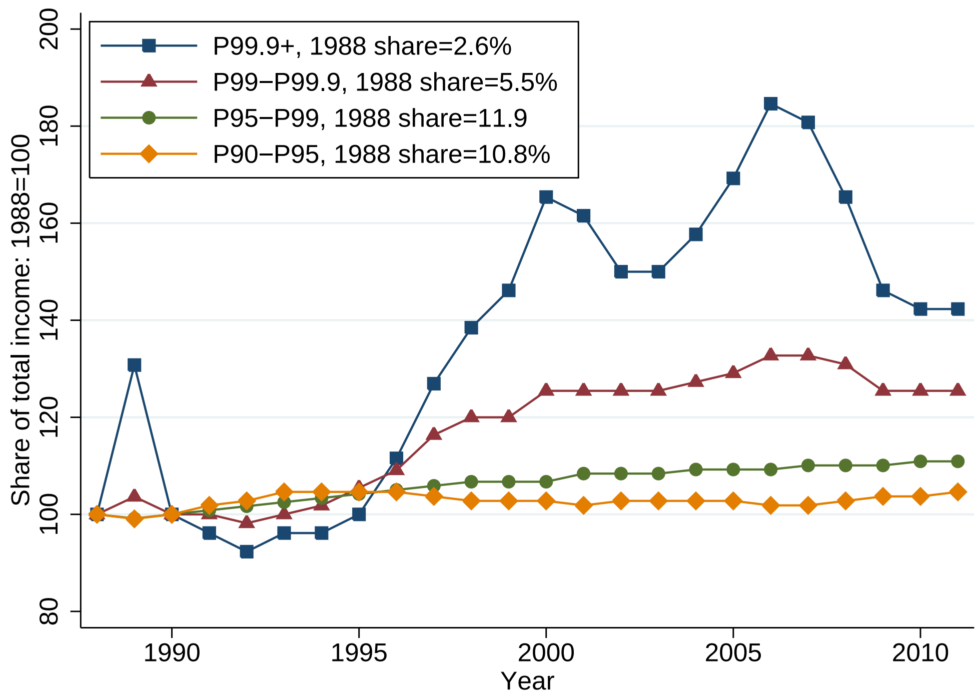
Source: CANSIM 204-0002. Shown is the share of income above the indicated cutoff percentile.

province. For some smaller provinces, higher thresholds were suppressed in the original data to maintain confidentiality.

The table reveals two important trends. First, the cutoff thresholds for reaching P99.9 grew much more than for P95 in all provinces. Nationally, the growth in the P95 threshold was 18.9 percent and at P99.9 it was 52.3 percent. This suggests much larger income growth at the very top, consistent with previous findings. The second important feature of Table 1 is the vast difference across provinces. As found in Veall (2012), the growth was much stronger in some provinces than others, with Alberta leading the pack. For the threshold to be in P99, the cutoff grew by 19.7 percent in Nova Scotia over this period, but by 75.1 percent in Alberta.

We now turn to the share of income earned by those over the percentile group thresholds seen above. To provide a sense of overall trends in high incomes in Canada, Figure 3 shows the national time series for several top income shares over the period 1988 to 2011. There is clear growth above the 90th, 95th, 99th,

Figure 4: Top income shares, 1988 to 2011



Source: CANSIM 204-0002. Shown is the share of income between the indicated cutoff percentiles.

and 99.9th percentile. The percentage growth was greater for the higher fractiles shown in Figure 3. In the top ten percent (above the 90th percentile cutoff; P90) the share of income grew 12.4 percent between 1988 and 2011, from 32.2 percent of total income to 36.2 percent. Above P95, the growth was stronger at 17.3 percent over this time period; for the top 0.1 percent (above P99.9), the total income share rises by 38.7 percent.

The strength of the income share movements at the top of the distribution come into clearer focus by further parsing the data. Figure 4 breaks down the top 10 percent into four pieces: between P90 and P95, between P95 and P99, between P99 and P99.9, and above P99.9. This figure also uses an index rather than the absolute shares in order to emphasize the percentage change over time. There is no upward movement between P90 and P95. For the share between P95 and P99, there is growth of about 10 percent from from 12.1 percent in 1982 to 13.4 in 2011. Again, the growth is stronger in the higher fractiles. For the first 9 tenths of the top one percent, the income share grows from 6.2 to 7.4 percent, for an increase of 19.4 percent. For the top one tenth of the top one percent, the income share increases by 38.7 percent by 2011, but peaked earlier in 2006 at 84.6 percent above 1988.

These patterns for Canada are consistent with those in the U.S. (and elsewhere) in two important ways. First, the rise in incomes is concentrated at the very top of the distribution. Not only are tax developments at the top of the income distribution important for redistribution, but also these taxpayers have access to substantial financial advice that may facilitate tax avoidance. Second, Saez and Veall (2005) show the source of incomes among those at the top has shifted substantially over the last half century from capital income toward earned income. All else equal, this change would tend to make income shifting or tax avoidance more difficult now than in earlier times.

To preview the core empirical relationship we will explore in the regressions, we plot in Figure 5 the relationship between the tax rates on top income and the top one percent income shares, using our data which vary at the province-year level. The data points near the y-axis of the figure are for Alberta in the 2000s. For the main cloud of data, there is a clear negative relationship between the tax rate and the top one percent share not only for the cloud as a whole, but also within provinces.

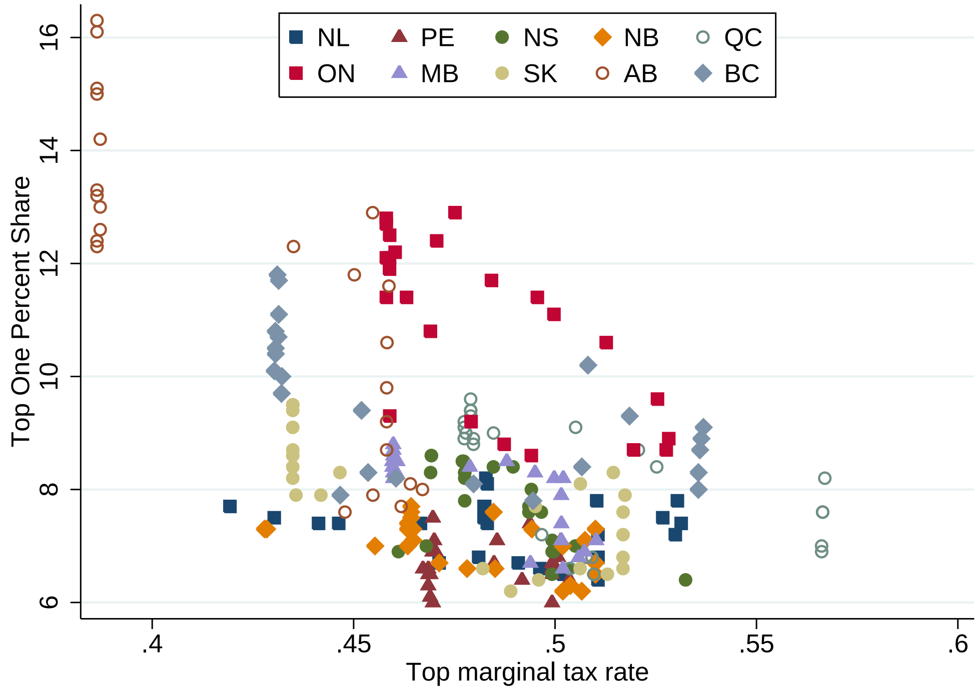
5 Main Results

In this section we present the main empirical results, then discuss the implications of our results. We begin by building up our specification, comparing the province-year results that can control for common time effects to time series estimates that do not. We then show the sensitivity of results to different specifications and samples, and how the results vary across different high income groups. We close the main results section with a discussion of the relevance of our results and an assessment of the contribution of taxes to the trends in high income concentration in Canada.

5.1 Empirical Results

The first results start with a time series specification like the ones used in the previous literature and then move toward the specification our data allow us to use in our context. These results appear in Table 2. The dependent variable in all cases is the log share ratio z^{P99} for the top one percent. The tax rate variable here is the log of one minus the marginal tax rate, calculated by using the actual average income

Figure 5: Top one percent tax rates and income shares, 1988 to 2011



Shown is the top tax rate for each province and the income share of the top one percent

of those in $P99$. We instrument for the actual marginal tax rate using the threshold marginal tax rate, as is common in the literature.[14] Here and in all specifications in this paper we weight by the number of taxfilers in the jurisdiction.[15]

For the national specifications in the first two columns, we use an average over the tax rates in all provinces, weighting each province in the average by total income. Using just the 24 annual national observations between 1988 and 2011 we obtain a strongly significant elasticity estimate of 2.046 in the first column of Table 2. Of course, this requires the assumption that there are no other time series trends

[14]The threshold marginal tax rate is calculated using the national average threshold income level for each fractile and each province's tax system. In practice for the fractiles we use, there is very little difference between using actual and threshold marginal tax rates because in Canada there is little progressivity at higher income levels. So, the same top tax rate applies to everyone in the top one percent fractile in almost every year. This is not true in many other countries, necessitating the commonly-used threshold instrumental variables approach.

[15]We base this choice on the discussion in Solon et al. (forthcoming). For a causal investigation like ours, weighting by population can be appropriate to account for heteroskedasticity. Solon et al. (forthcoming) note that it is not necessarily true that this correction can improve the precision of the estimation and recommend observing whether standard errors fall with weighting. In our case, the standard errors do fall with weighting, so we proceed to use the weights for our estimation.

Table 2: Comparing Time Series to Province-Year Results

	National Time Series		Province-Year Data		
	No Controls	Add US Top 1%	Province Controls	Add US top 1%	Add Time Controls
	(1)	(2)	(3)	(4)	(5)
Observations	24	24	240	240	240
R-Squared	0.473	0.864	0.710	0.882	0.941
Log (1-MTR)	2.046***	0.549*	1.673***	0.536*	1.068*
	[0.319]	[0.292]	[0.361]	[0.275]	[0.501]
Log US Top		1.009***		0.920***	
One Percent Share		[0.142]		[0.093]	

Note: The dependent variable is the log share ratio. MTR signifies the marginal tax rate. Province controls included in the last three columns. Estimation is by instrumental variables using threshold MTR as instrument. In brackets are robust standard errors, clustered by province for province-year data. One asterisk indicates significance at the 10 percent level, two asterisks for 5 percent, and three asterisks for 1 percent.

in top income shares and therefore attributes all of the changes in top income shares to taxes. In the second column, we follow Saez and Veall (2007) and include the log of the top one percent share from the United States as a proxy for global trends affecting top earners.[16] As found by Saez and Veall (2007), this one control knocks the coefficient down by 75 percent. The US coefficient itself is close to 1.00, which suggests a tight correspondence between trends in the US and in Canada. Other combinations of polynomial time controls (not shown) deliver highly variable results. Without an easy way to choose which set of time controls is the right specification, the range of estimates for the elasticity is quite high.

In the next three columns of Table 2, we use the province-year data. Here, we cluster our standard errors at the province level to account for correlated errors within the time series for each province. The first column includes just the set of province dummies α_p but omits the time controls λ_t seen in equation (7). The estimated elasticity is 1.673, not very different from the time series estimate in column (1). With the US top one percent included in column (4), the coefficient is again knocked down substantially. In column (5), we add the time effects λ_t. These control flexibly for any arbitrary differences across time periods that are common across the provinces. Here, the elasticity estimate is 1.068, significant at the 10 percent level. These results indicate the sensitivity of elasticity estimates in a time series context

[16]The source for these data is the top incomes database at the Paris School of Economics. http://topincomes.g-mond.parisschoolofeconomics.eu/

to the way time effects are controlled, and also the value of the province-year framework which can control for these time effects in a flexible way. On the other hand, the value is limited by the relevance of these estimates to provincial versus federal tax changes. In the presence of inter-provincial shifting, our empirical approach will estimate well the sensitivity of income to tax changes at the provincial level, but will overstate the responsiveness at the federal level.

We now try out some other variations on our specification in Table 3. In this table, we estimate forms of (7), with the log ratio dependent variable, time and province effects, and standard errors clustered on province. The basic specification in the first column is the same as in column (5) of Table 2. In the second column, we add a control for total income. Atkinson and Leigh (2010) control for GDP in their cross-country regressions to try to account for differential overall income growth in the different economies. Here in column (2), we use the log of total reported personal income within each province-year for our income control. The income control itself in column (2) has a significant coefficient of 0.729, suggesting a strong positive relationship between income growth and the share of income in the top one percent. The coefficient on the marginal tax rate falls, but remains strongly significant at 0.689. Our estimate here might be compared to the 0.62 elasticity coefficient that is reported in Department of Finance (2010) using a similar approach with Canadian data for the period 1995 to 2006. In results not reported in the table, we tried including the marginal tax rate at the 50th percentile to test the assumptions of our empirical approach in equations (2) and (3). With the 50th percentile marginal tax rate included, the coefficient on the 99th percentile tax rate remains almost unchanged at 0.689.[17] This suggests our choice to focus only on the top fractile tax rate in our main specifications is supported by the data.

The next two columns explore two other alternative specifications. In column (3), we follow the advice of Solon et al. (forthcoming) to check the results unweighted. The coefficient changes little, and of note the standard error for the weighted version is smaller, suggesting the weighting improved the efficiency of our estimates. The fourth column replaces our log ratio dependent variable z_{pt}^{P99} with the simpler log share s_{pt}^{P99}. Again, the results are little changed. In the final two columns of Table 3, we subject the specification from column (2) to stronger tests by including linear province-specific trends in column (5) and quadratic province-specific trends in column (6). Relative to column (2), the precision of

[17]The coefficient on the 50th percentile tax rate is 0.029 with a standard error of 0.169.

Table 3: Impact of Specifications on Elasticity Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Basic specification	Add income	Unweighted	Log share Dependent Variable	Linear Provincial Trends	Quadratic Provincial Trends
Observations	240	240	240	240	240	240
R-squared	0.941	0.97	0.949	0.97	0.975	0.988
Log (1-MTR)	1.068**	0.689***	0.723**	0.640***	0.794**	0.510*
	[0.441]	[0.238]	[0.293]	[0.210]	[0.323]	[0.264]
Log Total		0.729***	0.805***	0.631***	0.903***	1.523***
Income		[0.0767]	[0.0341]	[0.0690]	[0.179]	[0.209]

Note: The dependent variable is the log share ratio for the top one percent (except for column 4 with log share). Province and time dummies are included. MTR stands for marginal tax rate. Estimation is by instrumental variables using threshold MTR as instrument. In brackets are robust standard errors, clustered by province. One asterisk indicates significance at the 10 percent level, two asterisks for 5 percent, and three asterisks for 1 percent.

the estimates decreases, but the estimated elasticities stay in the same range at 0.794 for linear provincial trends, and 0.510 for quadratic trends. Not shown here, the results attenuate slightly as cubic provincial trends are added to 0.417 (0.245) and 0.379 (0.260) for quartic provincial trends. We view these results as quite robust. For the rest of our estimation, we use the specification in column (2). While the results do hold up to quadratic provincial trends, the robustness checks that follow do show more variability when we use the quadratic provincial trends specification.

Table 4: Different Time Periods

	Base	1988-1999	2001-2011	1995-2005
	(1)	(2)	(3)	(4)
Observations	240	120	110	110
R-squared	0.97	0.961	0.972	0.978
Log (1-MTR)	0.689***	0.948**	1.026	0.458**
	[0.238]	[0.390]	[0.829]	[0.179]
Log Total	0.729***	1.257***	0.481***	0.926***
Income	[0.0767]	[0.312]	[0.118]	[0.242]

Note: The dependent variable is the log share ratio for the top one percent. Province and time dummies are included. MTR stands for marginal tax rate. Estimation is by instrumental variables using threshold MTR as instrument. In brackets are robust standard errors, clustered by province. One asterisk indicates significance at the 10 percent level, two asterisks for 5 percent, and three asterisks for 1 percent.

We also check the results in different time periods in order to get a stronger sense of the sources of identification. The first column in Table 4 replicates our base specification for the P99 top one percent group. The next two columns break out two distinct time periods. The first is 1988 to 1999, which was the period between the 1988 tax reform and the ‘tax on income’ reform implemented by provinces over the years 2000-2001. As can be seen in Figure 2, there was a large degree of variation in top tax rates across provinces over this time period. The estimate in column (2) for 1988 to 1999 is a statistically significant 0.948. The second time period we choose is 2001 to 2011, which is entirely within the ‘tax on income’ era. Here, there is much less within-province variation evident in Figure 2. Alberta is constant over this time period, and other provinces (with the exception of Newfoundland and Labrador) don’t move very much. The point estimate for this time period in column (3) is 1.026, but the standard error is too large to allow for a statistical inference different from zero.

In the final column of Table 4, we examine what happens over a narrow time window from 1995 to 2005 surrounding the ‘tax on income’ reform. Here, the point estimate is a statistically significant 0.458. Given the sharpness of the reform in 2000-2001, it might seem natural to attempt a regression discontinuity analysis around this reform. In trying out regression discontinuity specifications, we found that the existing upward trend in the high income shares over this period moved in the same direction as the predicted tax effect. The estimated tax elasticity, therefore, was highly dependent on what specification was used for the polynomial in time to control for the ‘running’ variable. We take this as further evidence of the difficulty of using national-level shocks or experiments to identify elasticities in the presence of such strong prevailing trends.

Next, we want to see how sensitive the results are to different definitions of income. The norm in much of the elasticity of reported income literature (see Saez et al. 2012) is to use a measure of ‘market’ income (excluding government transfers), and also excluding capital gains. The focus on market income is to facilitate comparison across long time spans during which the taxation and reporting of government transfers may have changed. Since high earners have a fairly low share of their income coming from government transfers, this assumption is normally assumed to be innocuous. In our case, the time period we cover has no important changes to the definition of income regarding transfers, so we take the broader total income definition as our default. In addition, the focus on earnings can be motivated by

Table 5: Different Income Definitions

	Total Income		Market Income	
	No Capital Gains	With Capital Gains	No Capital Gains	With Capital Gains
	(1)	(2)	(3)	(4)
Observations	240	240	240	240
R-squared	0.97	0.96	0.964	0.953
Log (1-MTR)	0.689***	0.817**	0.723***	0.791**
	[0.238]	[0.364]	[0.243]	[0.335]
Log Total	0.729***	0.766***	0.605***	0.643***
Income	[0.0767]	[0.0990]	[0.0527]	[0.0735]

Note: The dependent variable is the log share ratio for the top one percent. Province and time dummies are included. MTR stands for marginal tax rate. Estimation is by instrumental variables using threshold MTR as instrument. In brackets are robust standard errors, clustered by province. One asterisk indicates significance at the 10 percent level, two asterisks for 5 percent, and three asterisks for 1 percent.

noting that it is earnings processes that we expect to be affected by taxation.

The exclusion of capital gains in the literature (again see Saez et al. 2012) is motivated by its differential taxation, the high degree of year-to-year variability in aggregate gains, and the particular dynamics of capital gains realization around tax reforms. Evidence on the sensitivity of high income trends to these definitional choices suggests they are not pivotal. Veall (2012) examines the capital gains issue, while Milligan (2013b) explores different market and after-tax income definitions using the Canadian census. Nevertheless, excluding capital gains from the dependent variable could in principle omit important avenues for taxpayer responses to tax reforms.

Table 5 displays the results for the four income definitions available in the CANSIM data. The first column is the definition we have used as our default, total income without capital gains. When capital gains are added, the elasticity in this top one percent sample increases from 0.689 to 0.817. However, some of this extra impact may be transitory, relating to the dynamics of capital gains realizations. In the third and fourth columns are the results using market income. The results are quite close to those using total income, so our choice to use total income as our default definition is not consequential to our findings.

We turn in Table 6 to comparing the elasticity across different fractiles of income, ranging from P90

Table 6: Elasticities across Income Fractiles

	P90	P95	P99	P99.9
	(1)	(2)	(3)	(4)
Observations	240	240	240	190
R-squared	0.962	0.969	0.970	0.952
Log (1-MTR)	0.0246	0.221	0.689***	1.451***
	[0.219]	[0.218]	[0.238]	[0.541]
Log Total	0.424***	0.511***	0.729***	0.893***
Income	[0.0533]	[0.0636]	[0.0767]	[0.162]

Note: The dependent variable is the log share ratio for the indicated fractile. Province and time dummies are included. MTR stands for marginal tax rate. Estimation is by instrumental variables using threshold MTR as instrument. In brackets are robust standard errors, clustered by province. One asterisk indicates significance at the 10 percent level, two asterisks for 5 percent, and three asterisks for 1 percent.

up to P99.9. The source data for P99.9 are suppressed for all years for Newfoundland and Labrador and for Prince Edward Island, so we cannot use those observations. As well, 1988 and 1989 are suppressed for New Brunswick. The first column shows our result for P90. The estimated elasticity is 0.025, which is small and not statistically significant. In the next column for P95, the point estimate is larger at 0.221, but still statistically insignificant. The third column replicates our base specification for the top one percent using P99. Finally, the fourth column reports an elasticity estimate of 1.451 for those in P99.9. This is more than twice the estimated elasticity for those in P99. Our results here are consistent with the international literature, which shows increasing responsiveness in the very top fractiles.

The final set of results in this section breaks down the top ten percent of incomes in a different way. We cut the top ten percent into four pieces: from P90 to P95; from P95 to P99; from P99 to P99.9; and from P99.9 up. This breakdown allows us some insight into how much of the results in Table 6 were driven by those at the very top. Since our empirical approach outlined earlier uses the complement to the high-income fractile as an implicit control group, we must exclude from the share calculation the income above the threshold. For example, for the P90 to P95 share we exclude the income earned in the top five percent, so that the P90 to P95 share represents how much of the bottom 95 percent of income is captured by those between P90 and P95. This provides further context for the results shown in Table 6 and how they vary across the top ten percent of the income distribution.

The interior fractile results appear in Table 7. In the first column the estimated elasticity is -0.214 but

Table 7: Elasticities for Interior Fractiles

	(1)	(2)	(3)	(4)
	P90-P95	P95-P99	P99-P99.9	P99.9+
Observations	240	240	190	190
R-squared	0.852	0.954	0.973	0.952
Log (1-MTR)	-0.214	-0.0773	0.364**	1.451***
	[0.157]	[0.157]	[0.150]	[0.541]
Log Total	0.120**	0.244***	0.549***	0.893***
Income	[0.0472]	[0.0426]	[0.0379]	[0.162]

Note: The dependent variable is the log share ratio for the indicated fractile. Province and time dummies are included. MTR stands for marginal tax rate. Estimation is by instrumental variables using threshold MTR as instrument. In brackets are robust standard errors, clustered by province. One asterisk indicates significance at the 10 percent level, two asterisks for 5 percent, and three asterisks for 1 percent.

statistically insignificant. In column (2), the point estimate for the range P95 to P99 is also negative and statistically indistinguishable from zero. The third column has the results for the bottom nine tenths of the top one percent, P99 to P99.9. Here, the estimate is a positive and significant 0.364. Finally, the top P99.9 percentile group shows an elasticity of 1.451, which is highly significant and large. These results suggest that the 0.689 elasticity we estimated in Table 6 for P99 is driven largely by those in the top tenth of one percent, where the elasticity is four times what is observed in the lower nine tenths.

5.2 Relevance of Results

We now place our estimates in context through two exercises. First, we combine our estimates with the standard formula for the revenue-maximizing tax rate and compare these results to the existing top rates in Canada. This allows one to observe the potential scope for provincial governments in Canada to increase tax rates at the top while still gaining revenue. The second exercise attempts to calculate how much of the observed trends in high income concentration can be explained by changes in top marginal tax rates.

The first exercise considers the scope for higher taxes in Canada, given our estimates. Our primary finding for the elasticity of reported income for the top one percent in group P99 is an estimate of 0.689, and for the top 0.1 percent in group P99.9 our preferred estimate is 1.451. The literature review in Saez et al. (2012) concludes that a reasonable range for the long-run elasticity is 0.12-0.4; Diamond and Saez

(2011) use 0.25 as the central estimate of the elasticity for their policy analysis. Department of Finance (2010) also surveys the literature, finding several studies with estimates over 0.6 for higher earners. In this context, our estimate of 0.689 for P99 is high, and 1.451 for P99.9 very high. Again, this estimate is most relevant for provincial tax changes, as inter-provincial shifting would change the analysis for federal level changes.

The magnitude of our estimates can be put into context by calculating the revenue-maximizing tax rate τ^*, which is the rate corresponding to the peak of the so-called 'Laffer Curve'. At this point, an incrementally higher rate will raise no further net revenue as the mechanical effect of the tax increase will be completely offset by the behavioural response of lower taxable income. This revenue-maximizing tax rate considers only the income tax revenue, and is not in general the optimal tax rate.

Saez et al. (2012) derive the formula for τ^* for a tax bracket affecting some proportion of higher earners

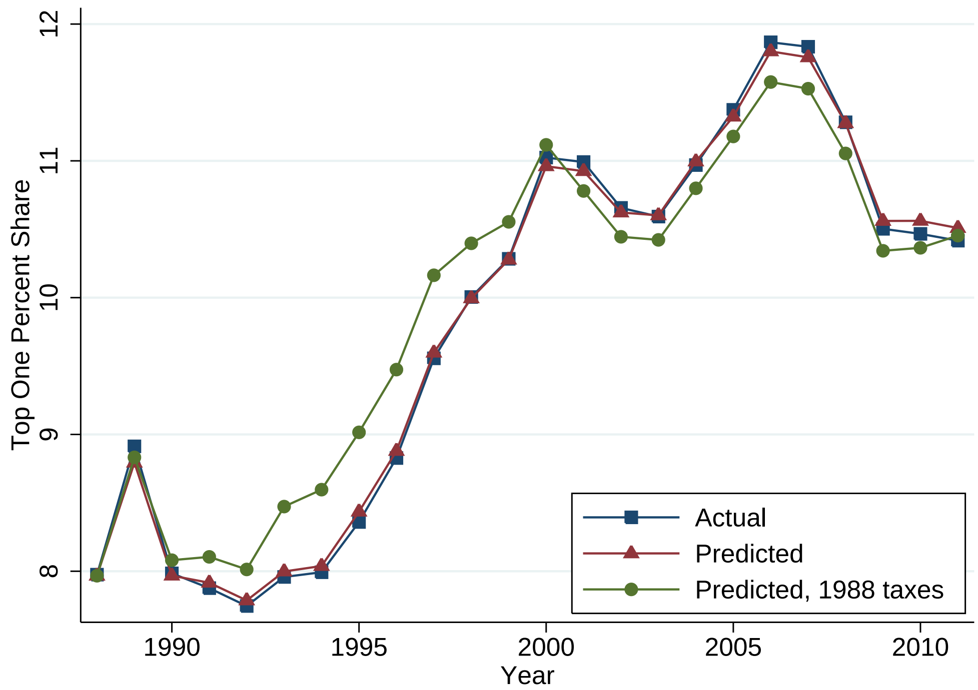
$$\tau^* = \frac{1}{(1 + a \times e)}, \tag{8}$$

where a is a parameter of the income distribution and e is the elasticity of reported income. For Canada, a reasonable estimate of a is 1.81.[18] Plugging $a = 1.81$ and $e = 0.689$ into equation (8) yields an estimate for τ^* of 44.4 percent. In Figure 1, four provinces have a top marginal tax rate for 2013 under 44.4 percent and six provinces are higher. Using the P99.9 estimate of 1.451, the revenue maximizing tax rate τ^* would be only 27.5 percent. If true, this would suggest all provinces could increase revenue by lowering the tax rate for those in income group P99.9. On the other hand, the estimated elasticity for P99-P99.9 in Table 7 is 0.364, which delivers a τ^* of 60.3 percent. This is well above the top tax rate in all provinces and suggests substantial room for higher taxes before hitting the revenue-maximizing rate for income levels up to P99.9.

We now turn to the second exercise, assessing the contribution of high-income tax rates on the observed trends in high income shares in Canada over the period from 1988 to 2011. The motivation for

[18]The parameter a is derived by taking the average income $\overline{y^g}$ over some threshold Y^g, and taking the ratio $\beta = \frac{\overline{y^g}}{\overline{y^g} - Y^g}$. For Canada, the Canada Revenue Agency Preliminary Statistics for the 2011 tax year report that the 203,010 Canadians with total incomes over \$250,000 have total assessed income of \$113,092,727,000. This yields an average of \$557,080. Using the formula for a, we find $a = 1.81$. Atkinson et al. (2011) reports in their Table 6 a value for β of 2.42 for Canada, which corresponds to an a for Canada of 1.704. This is fairly close to our estimate of 1.81.

Figure 6: Predicted top one percent share using 1988 tax rates



Shown is the actual top one percent national income share, as well as predicted values from the estimated model. Also shown is the predicted values using the 1988 top one percent marginal tax rate.

this exercise is to ascertain whether our estimates are large enough to account for much of the observed trends in high income concentration in Canada.

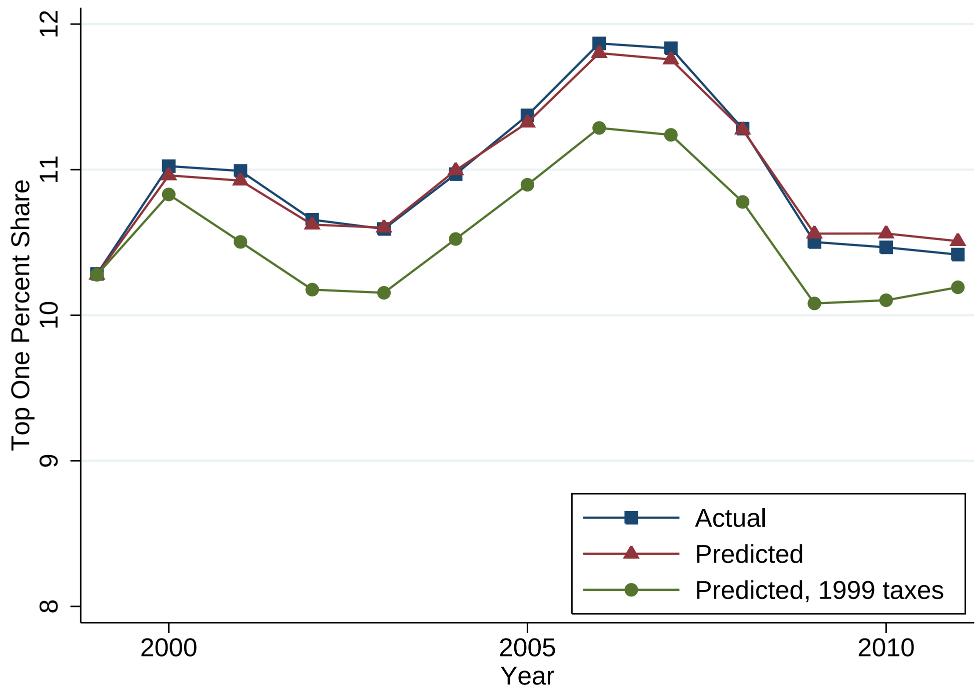
We begin by taking our preferred estimation from Table 3 column(2) and imposing the 1988 tax rates on all years in the data. We then use our estimates to predict values for each province and year with the counterfactual 1988 tax rates, then aggregate across the provinces within each year (weighting by total income) to produce a national time series. The predicted values show an estimate of what the P99 share would have been if tax rates had stayed at their 1988 level. The results are graphed in Figure 6.

There are three lines in Figure 6. The line indicated by the squares shows the actual series for the top one percent income share from 1988 to 2011. The line indicated by the triangles shows the path for the top one percent share predicted by our model. The lines are nearly on top of each other, indicating that our model does a very good job of fitting the trends. The third line indicated by circles shows the predicted values when we hold taxes level at their 1988 value. Through the 1990s, the predicted top one percent share would have been even higher with 1988 taxes, as most provinces raised their taxes over this period. (Figure 2 shows these provincial tax trends.) In the period since 2000, tax rates fell, meaning that the income share of the top one percent rose more than would have been the case if taxes had stayed constant. For this period, the predicted series with 1988 taxes is slightly under the actual series. However, the largest impression from Figure 6 is the close correspondence of the 1988 tax predicted series to the actual series, suggesting that taxes do little to explain the rise in the top income share. Veall (2012) argues that it is hard to make the case that taxes have driven the increase in top income shares, as the time series for taxes and income shares just don't line up. Our evidence concords with Veall's argument.

Since 2000, however, tax rates have fallen considerably. In 1999, the income-weighted average of the top tax rate across Canada was 49.5 percent; by 2009 this dropped to 44.9 percent. The bulk of this decrease happened between 1999 and 2001 with the introduction of the 'tax on income' system for the provinces and the cancellation of the high income surtax federally. Over the time period since 1999, therefore, our estimates may have some scope to explain the changes in the top one percent income share. We pursue this possibility by performing another simulation for the years since 1999, applying the 1999 tax rates to all years in this sample. As above, this allows us to predict a counterfactual path for the P99 income share in the absence of any tax change since 1999. We again note that our estimates are best applied at the provincial level and for federal tax changes may overstate the elasticity if there is substantial inter-province income shifting.

The results are graphed in Figure 7. The actual top share series rises from 10.3 to 11.9 between 1999 and 2006, for a gain of 15.4 percent. However, our simulation attributes 0.6 of the 1.6 percentage point gain to the change in marginal tax rates at the top. Put another way, our estimates suggest that the drop in marginal tax rates in the 2000s can account for 36.5 percent of the increase in the top income share over this time period. So, while the overall long-run trend toward higher income concentration cannot be explained by taxes, the drop in tax rates in the 2000s did contribute some to the income concentration seen during that decade.

Figure 7: Predicted top one percent share using 1999 tax rates



Shown is the actual top one percent national income share, as well as predicted values from the estimated model. Also shown is the predicted values using the 1999 top one percent marginal tax rate.

6 Extensions

In this final section of the paper, we present two extensions of our work. The first looks at some differences across provinces. The second extension uses another source of data to examine the most sensitive components of income to changes in top tax rates.

6.1 Provincial heterogeneity

There are several reasons why the results might differ by province. The income thresholds to be in the top one percent vary substantially across provinces, as can be seen in Table 1. If someone's ability to respond to higher taxes depends on the absolute level of income, this would make provinces with lower one-percent thresholds different than those with higher thresholds. In addition, Alberta's provincial top marginal tax rate of 10 percent is the lowest in Canada. When added to the federal tax rate of 29 percent, Alberta high income earners face a marginal tax rate of 39 percent, more than 10 percentage points lower than in some other provinces. This suggests some return for finding ways to shift income across provinces. As one example, assets transferred to a trust resident in Alberta are taxed using Alberta tax rates. This kind of arrangement would transfer taxable income out of a higher tax province and into Alberta. The other side of this coin is that residents of Alberta potentially have fewer instruments to avoid taxes, since they cannot shift income to a lower tax province. This suggests that to the extent that income shifting is important, the elasticity of reported income in Alberta might be lower. Of course, there are other institutional factors that differ in Alberta so this analysis can only be suggestive.

In Table 8, we begin by reproducing the base results using our regular specification. Here, we do not use the instrumental variables estimator because of the difficulties of doing so with the interaction terms we introduce in the other columns in the table. The second and third columns show the results for a specification in which we allow for a different response in provinces from Ontario to the west. The estimate of 0.619 in column (2) is for the five eastern-most provinces, while the five more western provinces see a significant 0.360 higher elasticity. This may reflect the higher average income thresholds to be in the top one percent in the western provinces. The same pattern holds for the top one tenth of one percent in column (3).

In the fourth column, we turn the focus to Alberta by interacting the marginal tax rate with the Alberta fixed effect in order to estimate a different impact of the marginal tax rate in Alberta. While there is little change to the main effect estimate at 0.619, the interaction on the Alberta term shows a point estimate of -0.504. While this is statistically insignificantly different from zero, the point estimate is nearly large enough to counteract the main effect leaving the predicted elasticity for Alberta close to zero. In the fifth column, we repeat the interaction specification for the P99.9 percentile group. Here, the main point estimate for the marginal tax rate is large at 1.492, but the interaction with the Alberta tax effect is -1.406, significant at the 10 percent level. While the standard errors in both columns 4 and 5 are large, the results are consistent with the elasticity in Alberta being lower. The results for the West and Alberta together indicate that the larger western elasticities from columns (2) and (3) are not driven by Alberta. This may indicate that geographic proximity to Alberta (the lowest-tax province) could play a role in income shifting opportunities. These interesting differences across provinces provide opportunities for

Table 8: Provincial Interactions

	(1)	(2)	(3)	(4)	(5)
	Base	West interaction	West interaction	Alberta interaction	Alberta interaction
	P99	P99	P99.9	P99	P99.9
Observations	240	240	190	240	190
R-squared	0.968	0.97	0.954	0.969	0.954
Log (1-MTR)	0.615*	0.619**	1.215*	0.619**	1.492*
	[0.275]	[0.270]	[0.578]	[0.270]	[0.637]
Provincial dummy		-1.612***	-1.361*	0.849***	0.802***
		[0.299]	[0.685]	[0.0579]	[0.140]
Provincial Dummy* log (1-MTR)		0.360*	0.464	-0.504	-1.406*
		[0.165]	[0.370]	[0.453]	[0.655]
Log Total Income	0.730***	0.863***	1.264***	0.863***	1.264***
	[0.0872]	[0.177]	[0.284]	[0.177]	[0.284]

Note: The dependent variable is the log share ratio for the indicated percentile group. Province and time dummies are included. MTR stands for marginal tax rate. Columns two and three interact a dummy for being in the five western provinces with $log(1-MTR)$. Columns four and five do the same, but with a dummy for Alberta. In brackets are robust standard errors, clustered by province. One asterisk indicates significance at the 10 percent level, two asterisks for 5 percent, and three asterisks for 1 percent.

continuing research.

6.2 Income components

In this set of results, we investigate the response of different components of income to the marginal tax rate among high earners. These results may shed some light on the mechanisms that underlie the overall tax elasticities found in this paper. We pursue this analysis using data drawn from the Canada Revenue Agency Income Statistics, which report a comprehensive set of disaggregated income and tax measures annually and by province, grouped into different income categories. These data were previously used by Gagné et al. (2010) to estimate tax elasticities, and by Saez and Veall (2007) to create top income share series.

Our data run from 1991 to 2010. The largest challenge to making use of these data is the nominal thresholds for the income categories. For example, the number of taxfilers and total amount for various definitions of income are reported for a category that includes taxfilers with total reported income of

$250,000 or more for each year in our sample. However, this $250,000 nominal threshold includes 0.19 percent of all taxfilers in Alberta for 1991, but 1.38 percent by 2010. When we observe that the total income share of this $250,000 group rose from 4.31 percent to 15.60 percent in Alberta from 1991 to 2010, the growth is comprised of two parts. Part of this growth is due to increasing income concentration, but part of it is just 'bracket creep' as more of those with constant real income are pushed over the nominal threshold. Since we are comparing across provinces where the $250,000 cutoff cuts at different points of the tail of the income distribution, this introduces unwelcome province-year varying noise to our estimate of top income shares.

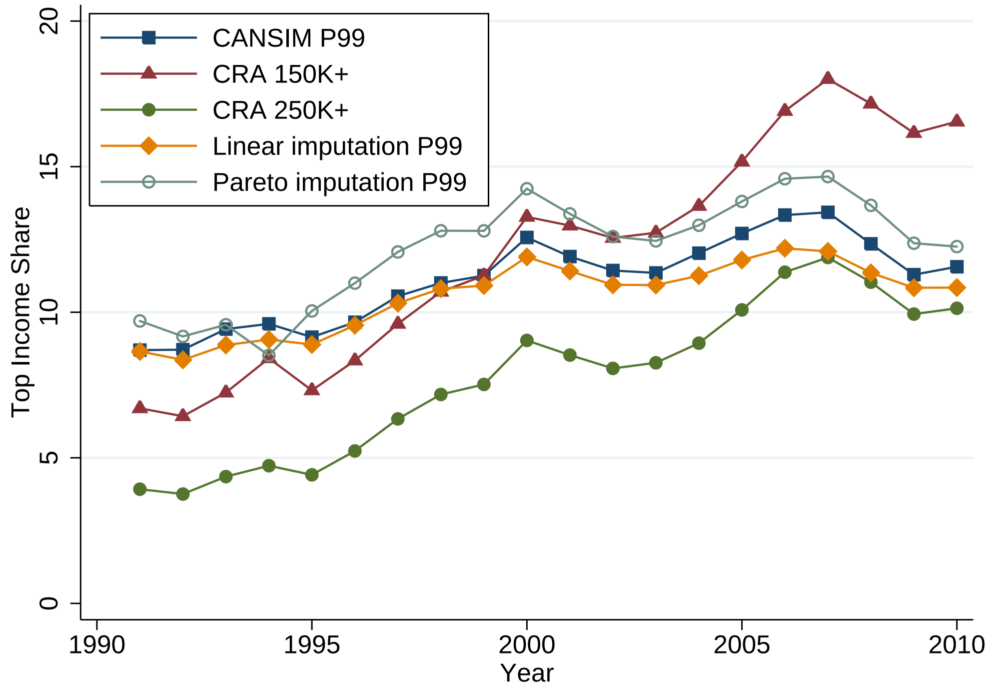
With these caveats, we present results showing the regression results for the top $250,000 category, and for those with incomes higher than $150,000. We also take two approaches to improve on these raw results. We use the Canada Revenue Agency (CRA) data to create our own top income fractile shares and cutoffs by interpolation. We try two methods for the interpolation. The first is a basic linear interpolation that assumes homogeneity of those within an income category. The second employs the Pareto distribution assumption to make the imputations, which allows for income concentration at the top of each category when performing the imputations. Each method is described in the data appendix.

A comparison of our constructed income share data to the CANSIM data and the raw CRA data appears in Figure 8. We aggregate across provinces to a national series using total income as weights. We also plot the share in the $250,000 category and the share above $150,000 from the CRA data.

As expected, the raw category data do not match the CANSIM data well. In the early 1990s, the $250,000 category had much less than one percent of taxfilers, but the proportion of taxpayers in this category grows through time. The $150,000 and higher category is a better match for the top one percent in the early years, but far exceeds the CANSIM numbers by 2010. Our constructed imputations fit the data fairly closely. The linear imputation fits very closely until around 2000, when it begins to underpredict the CANSIM series. This is expected, since this imputation assumes homogeneity within each category when in reality income within categories is likely skewed toward the top. The Pareto imputation, in contrast, overpredicts the income share of the top one percent compared to the CANSIM data.

We wish to use the CRA data to examine the components of income. Since the categories are defined based on total reported income, rather than the components we wish to study, we cannot use the exact

Figure 8: Top Income Shares using CRA data



Shown are four series derived from Canada Revenue Agency Income Statistics and the CANSIM P99 series.

same methodology as we do to impute total income. Instead, we follow this procedure. We define our interest as the income components of those in the top fractiles of total income. To find the component income share of a top fractile of total income, we apply the cumulative distribution function of total income to the observed category totals of the component. For example, if the top one percent of total income spreads across all of the $250,000 category and 40 percent of the $150,000 to $250,000 category, we take all of the component income from the $250,000 category and 40 percent of the component income from the $150,000 to $250,000 category.

In addition to breaking the data down into income components, we also report results for several tax base adjustments. These adjustments take the form of deductions from total income and non-refundable tax credits. Our results are reported in Table 9 for the $150,000+ shares and the $250,000+ shares, as well as the P99 and P99.9 results for both our linear and Pareto imputation methods.

The first row of Table 9 shows the results in this sample using the CANSIM data. The results are

Table 9: Components of Total Income in CRA data

	150+ Category	250+ Category	Linear P99	Pareto P99	Linear P99.9	Pareto P99.9
	(1)	(2)	(3)	(4)	(5)	(6)
CANSIM Total Income			0.562*	0.562*	1.041**	1.041**
			[0.322]	[0.322]	[0.482]	[0.482]
Total income	0.943***	1.321***	0.670***	0.983***	0.452***	1.559***
	[0.310]	[0.340]	[0.255]	[0.258]	[0.167]	[0.456]
Total Employment	0.929**	1.289***	0.522*	0.984***	0.45	1.534***
	[0.366]	[0.400]	[0.292]	[0.329]	[0.290]	[0.567]
Total Self-Employment	0.624	0.583	0.0421	0.351	-0.272	0.889
	[0.800]	[1.354]	[0.706]	[0.911]	[0.930]	[0.928]
Total Capital income	0.815	1.216***	0.689*	0.882**	0.189	1.396***
	[0.514]	[0.450]	[0.399]	[0.419]	[0.251]	[0.505]
Taxable Dividends	1.213*	1.850**	1.022	1.957**	0.734	2.110**
	[0.670]	[0.775]	[0.699]	[0.994]	[0.680]	[0.967]
Interest and Investment Income	0.241	0.469	0.0142	0.177	-0.38	0.69
	[0.386]	[0.449]	[0.504]	[0.453]	[0.386]	[0.513]
Capital Gains	-0.547	0.251	-0.756	-1.235	-0.653	0.685
	[0.968]	[0.386]	[0.979]	[1.777]	[0.643]	[0.714]
Taxable income	1.151***	1.507***	0.850***	1.196***	0.628***	1.734***
	[0.261]	[0.348]	[0.170]	[0.177]	[0.115]	[0.400]
Total Deductions	0.227	0.612	0.0719	0.204	-0.239	0.876
	[0.711]	[0.646]	[0.619]	[0.738]	[0.640]	[0.853]
Total non-refundable Credits	0.363	0.475	-0.283	0.201	-0.337**	0.72
	[0.353]	[0.478]	[0.276]	[0.251]	[0.169]	[0.458]
Charitable donations	0.505	0.576	0.363	0.367	-0.251	0.868*
	[0.463]	[0.604]	[0.307]	[0.464]	[0.380]	[0.518]
RPP contributions	-0.00692	1.740***	-0.737	0.517	0.879	2.135**
	[0.498]	[0.648]	[0.541]	[0.653]	[0.864]	[0.936]
RRSP contributions	0.633	0.47	0.203	0.533**	-0.375**	0.735*
	[0.395]	[0.362]	[0.136]	[0.247]	[0.176]	[0.385]
Union / Professional Dues	0.719	2.577**	0.161	1.661***	1.413**	2.720***
	[0.874]	[1.095]	[0.378]	[0.437]	[0.660]	[0.645]
Personal amounts	0.342	0.685**	-0.226	0.394**	-0.133	0.959***
	[0.235]	[0.318]	[0.174]	[0.175]	[0.114]	[0.333]

Note: These results use the CRA income statistics data for years 1991-2010. (Except the first row which has the CANSIM data for comparison.) The dependent variable is the log share ratio for the indicated measure. Province and time dummies are included, as well as the log of total income. In brackets are robust standard errors, clustered by province. One asterisk indicates significance at the 10 percent level, two asterisks for 5 percent, and three asterisks for 1 percent.

comparable to our base results in Table 6, although point estimates are smaller. The regression results using the CRA data make up the rest of the rows in Table 9. We use our instrumental variables approach for the estimation here, with the log share ratio dependent variable. The first two columns of the table display the results for the $150,000+ and the $250,000+ category. The results show a strong sensitivity of Total Income to the tax rate, with strongly significant estimates. There are especially strong effects for Total Employment income, Total Capital Income, and Taxable Dividends. Of course, given the nominal thresholds, these results are difficult to interpret.

The next two columns have our estimates for the P99 group using the linear and the Pareto imputation, respectively. In almost all cases, the Pareto estimates show more sensitivity to the tax rate than the linear estimates. With the CRA data using the linear method, the estimated elasticity for total income is 0.670, quite similar to our earlier results with the CANSIM data. Using the Pareto imputation the estimated elasticity for total income becomes 0.983, larger than our earlier results. The value of the CRA data comes from our ability to also look at the components of income. The definitions for the components are provided in the data appendix. There are strong responses by employment income and capital income, with dividend income in particular showing strong sensitivity. It is perhaps surprising that self-employment income doesn't display the same responsiveness. Taxable income for all specifications shows a stronger sensitivity than total income, which is consistent with expectations since the availability of deductions from total income increases the potential scope for tax shifting or avoidance. In the last two columns we present the results for the top 0.1 percent. Here, the Pareto estimates are much larger than the estimates using the linear imputation. This might be expected because the linear imputation is likely to do poorly picking out the top 0.1 percent from the top $250,000 category.

The bottom half of Table 9 reports deduction and credit items. Overall, total deductions don't show a significant effect. The only item that shows a consistently strong result is union and professional dues. Charitable donations become significant only for the top 0.1 percent in the last column. The lack of consistency of the estimates suggests that no strong conclusions can be drawn from the credit and deductions items.

7 Concluding remarks

In this paper, we use subnational variation in tax rates in Canada to estimate the elasticity of reported income for high earners. We find that the elasticity is large, but concentrated at the very top of the income distribution. Although the responsiveness is large, the thirty-year surge in high income concentration is largely unrelated to tax rates. We also find that there are important differences when comparing other provinces with Alberta, and also the response is present for both earned and capital income. These findings suggest multiple channels for income shifting and tax avoidance that may underlie the estimates we uncover.

An important limitation of our work is its applicability to federal tax rate changes. Since we use provincial-level variation, our estimates embody any inter-provincial income shifting that might occur. Federal tax changes should not induce inter-provincial income shifting, so our estimates are best used to predict the impact of provincial-level tax policy changes.

Another limitation is our treatment of dynamics, as we assume a contemporaneous reaction between tax rates and income. So, our estimate is clearly a short-run elasticity. As discussed by Saez, Slemrod, and Giertz (p. 13-14, 2012) the long-run elasticity may be higher because the generation of income shifting opportunities may take time, but also might be lower because the short-run elasticity captures short-run temporal income shifting dynamics (such as the timing of bonuses or realization of dividends or capital gains) that are harder to avoid over longer time spans.

In ongoing work in Milligan and Smart (2013), we look more precisely at the case of Alberta to examine the role of an internal 'tax haven'. The federal dynamics are made more interesting by the use of the common federal tax base across most of the country, as a dollar shifted in response to one province's taxes imposes a vertical externality on federal finances. We believe the work presented here in this paper provides a firm starting point for further research on these types of tax interactions in a federation like Canada.

A Data appendix

The Canada Revenue Agency (and its predecessor agencies) publishes an annual report on the incomes of Canadians drawn from the universe of tax forms that are filed. The data are available nationally and by province. The report includes data for income categories (e.g. total income, employment income, taxable income), deduction items (e.g. Registered Pension Plans, childcare expenses) and non-refundable credits (e.g. charitable donations and gifts, the basic amount, spousal amount). For each of these, the total amount and the number of taxfilers with any value for the item are listed. The data are presented in aggregate, as well as for income categories that are fairly constant across years.

We obtained the data for 1991 to 2010 by downloading the available years from the CRA website, For earlier years, we made a special request for a tabulation from the Canada Revenue Agency. The top category reported is for $250,000 for all years. Since this nominal limit doesn't change through time, 'bracket creep' means that more and more individuals overtake the nominal $250,000 threshold. To overcome this, we developed two methods to impute a consistent series, a linear interpolation and a Pareto interpolation.

The linear interpolation simply assumes everyone within an income category is homogenous. So, for the 2010 $250,000 category in Alberta, we observe that 1.38 percent of taxfilers are in the category with 15.60 percent of total reported income. The linear interpolation calculates the P99 share as $\frac{1}{1.38} \times (15.60) =$ 11.30 percent. Because people within the category are not homogenous but instead the income is concentrated closer to the top, this method underestimates top income shares. The comparable CANSIM value for the P99 share for 2010 in Alberta is 13.8 percent.

The Pareto interpolation makes the assumption that the upper tail of the income distribution is Pareto distributed. Pareto distributed variables have the property that the ratio of the mean above some cutoff to the value of the cutoff is a constant. In the CRA data, we observe the number of taxfilers and total income in the $250,000 category, so we can calculate the mean and divide by $250,000 to recover an estimate of this Pareto parameter. We then use this parameter to project the share. For example, in Alberta in 2010, there are 37,710 taxfilers in the $250,000 category who report $22,679,668,000 in total income, which is an average of $601,423. Divided by the threshold of $250,000 gives an estimate of 2.41 for the Pareto parameter. Using this, we estimate the Alberta P99 share to be 14.74 percent for 2010. This compares to 13.8 percent in the CANSIM data.

References

Atkinson, Anthony B, and Andrew Leigh (2010) 'The distribution of top incomes in five Anglo-Saxon countries over the twentieth century.' Technical Report, Australian National University, Centre for Economic Policy Research

Atkinson, Anthony B, Thomas Piketty, and Emmanuel Saez (2011) 'Top incomes in the long run of history.' *Journal of Economic Literature* 49(1), 3–71

Department of Finance (2010) 'The response of individuals to changes in marginal income tax rates.' Part 2 in *Tax Expenditures and Evaluations 2010*

Diamond, Peter, and Emmanuel Saez (2011) 'The case for a progressive tax: From basic research to policy recommendations.' *Journal of Economic Perspectives* 25(4), 165–190

Feldstein, Martin (1995) 'Effect of marginal tax rates on taxable income: A panel study of the 1986 Tax Reform Act.' *Journal of Political Economy* 103(3), 551–572

Fullerton, Don (1996) 'Comment on 'high-income families and the tax changes of the 1980s: The anatomy of a behavioral response' by Joel Slemrod.' In *Empirical Foundations of Household Taxation,* ed. Martin Feldstein and James Poterba (University of Chicago Press) pp. 189–192

Gagné, Robert, Jean-François Nadeau, and François Vaillancourt (2010) 'Taxpayers' response to tax rate changes: A Canadian panel study.' Scientific Series, CIRANO, December. Number 2000s-59

Milligan, Kevin (2013a) *Canadian Tax and Credit Simulator.* Version 2013-1

_ (2013b) 'Income inequality and income taxation in Canada: Trends in the census 1980-2005.' *University of Calgary School of Public Policy Research Papers* 6(24), 1–29

Milligan, Kevin, and Michael Smart (2013) 'The devolution of the revolution: Taxation of high incomes in a federation.' *Manuscript, Department of Economics, University of Toronto*

Piketty, Thomas, and Emmanuel Saez (2003) 'Income inequality in the United States 1913-1998.' *Quarterly Journal of Economics* 118(1), 1–39

Saez, Emmanuel (2004) 'Reported incomes and marginal tax rates, 1960-2000: evidence and policy implications.' In 'Tax Policy and the Economy, Volume 18' (MIT Press) pp. 117–174

Saez, Emmanuel, and Michael R Veall (2005) 'The evolution of high incomes in Northern America: Lessons from Canadian evidence.' *The American Economic Review* 95(3), 831–849

__ (2007) 'The evolution of high incomes in Canada, 1920-2000.' In *Top incomes over the 20th century,* ed. A B Atkinson and T Piketty (Oxford University Press) chapter 6

Saez, Emmanuel, Joel Slemrod, and Seth 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

Sillamaa, Mary-Anne, and Michael R Veall (2001) 'The effect of marginal tax rates on taxable income: A panel study of the 1988 tax flattening in Canada.' *Journal of Public Economics* 80(3), 341–356

Slemrod, Joel, and Wojciech Kopczuk (2002) 'The optimal elasticity of taxable income.' *Journal of Public Economics* 84(1), 91–112

Solon, Gary, Steven J Haider, and Jeffrey Wooldridge (forthcoming) 'What are we weighting for?' *Journal of Human Resources*

Veall, Michael R (2012) 'Top income shares in Canada: recent trends and policy implications.' *Canadian Journal of Economics* 45(4), 1247–1272