The Devolution of the Revolution: Taxation of High Incomes in a Federation

Kevin Milligan[*]
and
Michael Smart[†]

September 8, 2013
Preliminary Version

Abstract

We develop a theory of cross-border income shifting in response to personal taxation, and examine its implications for the revenue potential and excess burden of personal taxes at the subnational level. We estimate the elasticity of tax avoidance and of cross-border tax base shifting using data on top income shares for Canadian provinces around a significant reform in subnational taxation in Canada. According to our estimates, shifting taxable income between provinces accounts for about two-thirds of total tax avoidance in response to unilateral provincial tax changes. Implications for design of federal income tax policies are discussed.

1 Introduction

Cross-border tax avoidance has recently attracted increased attention from economists and policymakers. High tax rates on high-income individuals may induce them to migrate to lower-tax jurisdictions, or simply to shift assets and taxable income there, limiting the potential for such taxes to raise revenue and redistribute income efficiently. While popular attention to these issues has typically focussed on activities in the more notorious offshore "tax haven" countries, it is apparent that tax avoidance of similar magnitude may be occurring between jurisdictions within economic federations such as the EU and US, where mobility costs are low, and common institutions facilitate asset transfers.

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

Tax shifting within federations raises special issues, inasmuch as tax changes in one jurisdiction create fiscal externalities for others in the federation. Analytically, the revenue and welfare implications of tax avoidance may be measured by the elasticity of taxable income (ETI) with respect to the marginal tax rate (Feldstein, 1999). When there is a potential for cross-border tax avoidance, the ETI for unilateral tax changes by a single jurisdiction exceeds the ETI for coordinated tax changes by all jurisdictions, due to the effects of cross-border tax shifting.

In this paper, we extend the standard model of tax avoidance to incorporate tax shifting within a federation and show how the ETI may be decomposed into inter-jurisdictional income shifting and other forms of tax avoidance. We derive the implications of the model for the revenue potential and welfare cost of income taxation at the subnational and national levels. We then estimate the model using data on high-income shares for the Canadian provinces in the 1988-2010 period. Our data straddle the date of a major tax reduction by one province that anecdotal evidence suggests has created an "internal tax haven," with substantially greater incentives for tax shifting between provinces. In the Canadian case, we find that cross-province tax shifting comprises about two-thirds of the total ETI to unilateral provincial tax changes, suggesting that fiscal externalities are large.

The relatively small previous literature on cross-border effects of personal income taxation has generally examined specific mechanisms for tax shifting, rather than its aggregate effects on the ETI. In this vein, Kleven Landais and Saez (2012) find substantial effects of high-income tax rates on migration of "superstars" in Europe. These results complement and reinforce the earlier conclusions of Feldstein and Wrobel (1998) that state income tax differences lead to offsetting differences in equilibrium pre-tax wages, leaving little scope for redistribution at the subnational level. (In contrast, however, Young and Varner (2011) find little migration in response to a particular state tax reform in the US.) A parallel literature looks for evidence of mobility of financial assets in response to personal tax differences. The recent European Savings Directive in particular appears to have substantially reduces the use of offshore bank accounts, at least in some EU member countries (Johannesen, 2010).

There is now an extensive literature estimating the ETI for personal income tax systems, but most of it uses national-level data and ignores cross-border shifting. There are a few exceptions however. Gruber and Saez (2002) estimate an ETIs for the US that are larger for state than federal tax changes (though insignificantly so). However, they do not look for direct evidence of shifting across states. Giertz (2007) analyzes these issues in greater detail. Atkinson and Leigh (2010) estimate the ETI for five Anglo-Saxon countries using top income share data. Like us, they exploit differences in the timing of tax reforms across countries and employ a “difference-in-difference” estimator to control for unobservable trends in income distribution that are common to all countries. However, they do not investigate cross-border income shifting in response to tax reforms.

A few papers have received relatively little attention in the Canadian empirical literature. Mintz and Smart (2004) document the potential for shifting corporate income between provinces and estimate tax base elasticities, but they do not consider personal income shifting. Day and Winer (2006) examine policy-induced migration between Canadian provinces, but with a focus on the impacts of social insurance programs, rather than on tax rates faced by high-income taxpayers. Saez and Veall (2005) estimate the ETI at the national level using data similar to ours, but ignore provincial variation in tax rates. Veall (2012) documents the recent changes in top income shares at the provincial but does not estimate the effects of provincial taxes.

The plan of the paper is as follows. Section 2 introduces our model of tax avoidance and tax shifting, discusses its implications for the revenue and welfare cost of tax increases at the subnational level, and describes our strategy for estimating the key elasticities. Section 3 discusses personal income taxation in the Canadian federal system, the nature of the tax reform we study, and the potential for interprovincial tax shifting. Section 4 introduces our data, and Section 5 presents estimation results. Section 6 discusses policy implications and Section 7 concludes.

2 A model of tax avoidance and tax shifting in a federal system

A representative taxpayer in province i has potential income z and faces statutory marginal tax rate τ_i on taxable income above a fixed bracket threshold k. We note that the marginal tax rate τ_i comprises both a federal component T that is common to all provinces, and a provincial component $\tau_i - T$ that may vary among provinces. The taxpayer can shelter a dollars in income from tax by engaging in a tax avoidance activity, and can also shift s dollars of income to a low-tax province where the tax rate is $\tau_0 < \tau_i$. The taxpayer therefore declares taxable income

$$y = z - a - s$$

in the home province i, and pays tax $\tau_0 s$ on income s shifted to the low-tax province.

The private cost of tax avoidance is $C_a(a)$ and of interprovincial tax shifting is $C_s(s)$; both functions are strictly convex. The taxpayer then chooses (a, s) to maximize utility

$$u_i(\tau_i,\tau_0) = \max_{(a,s)} (1-\tau_i)z + \tau_i a + (\tau_i - \tau_0)s - C_a(a) - C_s(s) + (\tau_i - \bar{\tau}_i)k \qquad (1)$$

where $\bar{\tau}_i$ is the average tax rate on taxable income below the top bracket threshold k, under the tax system of province i. (Thus our model can be applied to any non-linear income tax system without loss of generality, as long as there exists an income level, or bracket threshold, above which the marginal tax rate is constant in all provinces.)

The first order conditions for (1) imply optimal avoidance and shifting activities are functions of tax rates of the form

$$a^* = a(\tau_i)$$

$$s^* = s(\tau_i - \tau_0)$$

That is, avoidance depends only on tax rate in the home province, and shifting depends only on the difference in tax rates between the home and low-tax provinces.[1]

From a policy perspective, the essential issue is the degree to which national and provincial tax bases respond to increases in marginal tax rates at home and in the low-tax province. In particular, we may define the semi-elasticity of avoidance activity as the percentage change in tax bases in response to a coordinated increase in all tax rates, or

$$e_a = -\left.\frac{\partial \log y}{\partial \tau_i}\right|_{(\tau_i - \tau_0)\text{ fixed}} = \frac{a'(\tau_i)}{y_i} \qquad (2)$$

Likewise, the semi-elasticity of shifting activity is the percentage change in tax bases in response to a decrease in the interprovincial tax differential alone, or

$$e_s = \left.\frac{\partial \log y}{\partial \tau_0}\right|_{\tau_i\text{ fixed}} = \frac{s'(\tau_i - \tau_0)}{y_i} \qquad (3)$$

The combined effect of a unilateral tax increase by province i on the tax base in province i is then the sum of the avoidance effect and the shifting effect, i.e.

$$e_u = \left.\frac{\partial \log y}{\partial \tau_i}\right|_{\tau_0\text{ fixed}} = -(e_a + e_s)$$

Understanding the magnitudes of these responses is evidently the key to evaluating the impact of tax rate changes on the revenues of federal and provincial governments, and on the excess burden of the tax system. Indeed, in the language of Chetty (2009), (e_u, e_s) are "sufficient statistics" for welfare analysis in this model.

[1] This is a consequence of additivity of net income in (a, s).

2.1 Estimation

To derive an estimating equation for these parameters, we may additionally assume that the private cost functions (C_a, C_s) are quadratic functions of avoidance and shifting activities, and are proportional to potential income,[2], i.e.

$$C_a(a) = \frac{1}{2e_a}\frac{a^2}{z} \qquad C_s(s) = \frac{1}{2e_s}\frac{s^2}{z}$$

The first order conditions for (1) then imply optimal avoidance and shifting functions

$$a^* = e_a \tau_i z_i$$
$$s^* = e_s(\tau_i - \tau_0) z_i$$

Then taxable income is

$$y = z[1 - e_u \tau_i + e_s \tau_0]$$

This function is non-linear in parameters (e_u, e_s), which is something of an impediment to estimation. We can linearize the tax base function by taking a first order approximation at the vector $\tau = \mathbf{0}$ to get

$$\log y \approx \log z - e_u \tau_i + e_s \tau_0 \tag{4}$$

Estimation of the model (4) comes down to specifying the structure of unobservables that are in log potential income $\log z$. We have a panel of taxable incomes and marginal tax rates for taxpayers in 10 provinces and 23 years. As discussed below, there have been substantial changes in the top tax rates of some provinces over the sample period, while tax rates of lower-income taxpayers have been more stable over time. With this structure of the data, we adopt two key empirical approaches to controlling for unobservable components of taxable income that may in principle be correlated with tax rate changes.

The first approach is the "share analysis" approach, common in the empirical literature on taxable income elasticities since Slemrod (1996).[3] We assume that taxpayers in each province i and year t may be divided into a group of high-income taxpayers $b = H$ who are affected by changes in top marginal tax rates, and a group of other taxpayers $b = L$ whose behaviour is unaffected by changes in tax rates, but who are affected by unobservable shocks to income in the same way on average as group H in the same province and year. Thus

$$\log z_{bit} = \alpha_{it} + u_{bit}$$

[2] These are common assumptions in the literature on international business tax avoidance; see e.g. Mintz and Smart (2004).

[3] See Saez, Slemrod and Giertz (2010) for a survey.

where α_{it} is the province–year-specific common shock, and u_{bit} is a stochastic term. Aggregating (4) across taxpayers in the two groups and differencing, then expressing the left-hand side of the resulting estimating equation in share form, yields

$$\log(\mu_{Hit}/(1-\mu_{Hit})) = -e_u\tau_{Hit} + e_s\tau_{H0t} + \Delta u_{it} \tag{5}$$

where μ_{Hit} is the share of high-income taxpayers in aggregate taxable income for province i and year t.

The second, related approach is to exploit the panel structure of our data. One problem with the share analysis approach is that it is not robust to the presence of unobservable changes in the distribution of income that are unrelated to, but contemporaneous with, reforms to the top marginal tax rate. Because our data permit us to compute top taxable income shares for taxpayers in each province separately, we can also include year effects in the regressions in order to control for unobservable changes in the distribution of income that are common to all provinces. To our knowledge, this has not been done before with subnational taxable income data (although see Atkinson and Leigh (2010) for a related approach to cross-country data).

Of course, the tax rate of the low-tax province τ_0 is common to all provinces in (5) and therefore collinear with year fixed effects. So a general difference-in-difference estimator is inadmissible when τ_0 is included in the regression. However, we do include year effects in a more standard specification of (5), in which only the home-province tax rate is used to explain changes in income shares. This allows us to compare our results to those of the previous literature, and to explore the robustness of results to inclusion of year fixed effects. In addition, we explore how results change when we replace the year effects with parametric controls for macroeconomic variables that proxy for non-tax factors influencing the distribution of income.

3 Personal taxation in Canada

Canada is a federation in which income taxation powers are co-occupied by the federal government and the governments of the ten provinces. Constitutionally, the provinces have wide latitude in designing their personal income tax systems, and they collect substantial revenue from them. In 2011, provincial personal income tax revenues were $73.8 billion, or 38% of combined federal–provincial revenues. Provinces generally apply progressive rate structures to taxable incomes, with top marginal tax rates currently (2013) ranging from 10% to 21% in the various provinces. In all provinces except Quebec, tax rates are applied to a common (federal) definition of taxable income, and taxes are collected on behalf of provinces by federal tax authorities.

A major reform to subnational taxation occurred in 2000/2001. Previous to this reform, provinces (outside Quebec) set their income taxes as a faction of "basic federal tax".[4] These rates ranged in 1995 from 69 percent in Newfoundland and Labrador to 45.5 percent in Alberta. An increase in this provincial tax rate affected all taxpayers proportionately. Provinces at this time had the ability to add income surtaxes for high earners in order to further manipulate the tax liability and marginal tax rates of those at the high end of the income distribution. A reform from this "tax-on-tax" to a "tax-on-income" system was implemented over 2000-2001. Under the new system, provinces could set their own brackets and rates, given the federally-determined taxable income.

While some provinces implemented the new tax-on-income system by choosing tax brackets and rates that produced tax liabilities identical (or very close) to the previous tax-on-tax system, the province of Alberta did not. Instead, Alberta adopted a flat-rate income tax with a top marginal rate of 10%, eschewing brackets altogether. The top rates implemented by other provinces ranged between 16 and 20%. These tax differentials appear to have led to new strategies for shifting taxable income to Alberta.

One strategy, widely promoted by tax advisers,[5] was for high-income taxpayers in other provinces to transfer personal assets to an inter vivos trust resident in Alberta; income received by the trust is taxed at the lower tax rate. The scheme appears especially popular in cases where a closely held corporation has substantial undistributed earnings or unrealized capital gains, so that tax savings exceed transaction costs of the trust. A trust is deemed resident in Alberta if a majority of its trustees reside there. In the words of one tax adviser marketing the scheme, "if the taxpayer does not know any Albertan residents they can choose an Alberta law firm or financial institution to act as trustee."[6]

For years, these trust arrangements attracted little notice from federal tax authorities, perhaps because provincial residency affects provincial but not federal tax revenues. In 2010, however, following a critical report from the federal Office of the Auditor-General, the Canada Revenue Agency announced a new

[4]The "basic federal tax" was the tax liability generated by the basic tax rate and tax bracket calculation. Basic Federal Tax excludes special surtaxes and abatements. Quebec had its own tax base, bracket, and rate structure.

[5]See, e.g., Martin Rochwerg, 2004, "Recent developments in estate planning: The Alberta advantage when using trusts," mimeo, Miller Thomson LLP, and Tim Cestnick, "Consider an Alberta trust even if you don't live there," *The Globe and Mail* (Toronto), June 21, 2003.

[6]Quoted from Cunningham Chartered Accountants, "Alberta and the use of Alberta trusts – the next tax haven?"

initiative to verify the residency of Alberta trusts and to scrutinize certain distributions from trusts to beneficiaries. As well, the province of Quebec, which administers an independent income tax law, requires resident taxpayers to self-report income received from non-resident trusts and to pay tax on it at home. Most recently, a ruling of the Supreme Court of Canada[7] has tightened rules on residency of trusts in a way that seems likely to severely restrict the future use of Alberta trusts for interprovincial tax avoidance. These recent developments are however unlikely to affect behaviour during the 1988-2010 period covered by our data set, described below.

While Alberta trusts appear to have been extensively used by to shift capital income to Alberta, they are practical only for the highest-income taxpayers, and they were not in themselves effective for shifting employment income out of high-tax provinces. An alternative tax planning strategy is simply for the taxpayer to declare residency in Alberta. Income taxes in Canada are payable in the province of residence of a taxpayer, irrespective of the location of employment. Residency for tax purposes is determined based on the taxpayer's principal residence on December 31 of each year. Moreover, federal tax authorities may not closely scrutinize provincial residency claims.[8] This situation may be contrasted to that of the US states, where nexus for individual income taxation typically reflects the location of employment as well as residence, and state tax authorities may aggressively pursue false claims of residency.

In short, the Canadian federal system is a useful testing ground for our theory of interjurisdictional income shifting in response to personal taxation. The anecdotal evidence points to the province of Alberta as a likely "onshore tax haven" towards which most domestic tax planning strategies were directed, at least in the period covered by our data set. Our empirical strategy is therefore to estimate our model with data on high-income shares in the other nine provinces, with marginal tax rates in that province and in Alberta as the principal explanatory variables of interest.

4 Data

Our strategy involves regressing reported income shares on tax rates and control variables. Below we first describe the source and formation of our income variables. Following that we outline how we calculate the tax rates. Finally, we provide information on our control variables.

[7] Fundy Settlement v. Canada, 2012 SCC 14

[8] See, e.g., "High-income earners use Alberta to save on taxes," *Calgary Herald*, April 30, 2013.

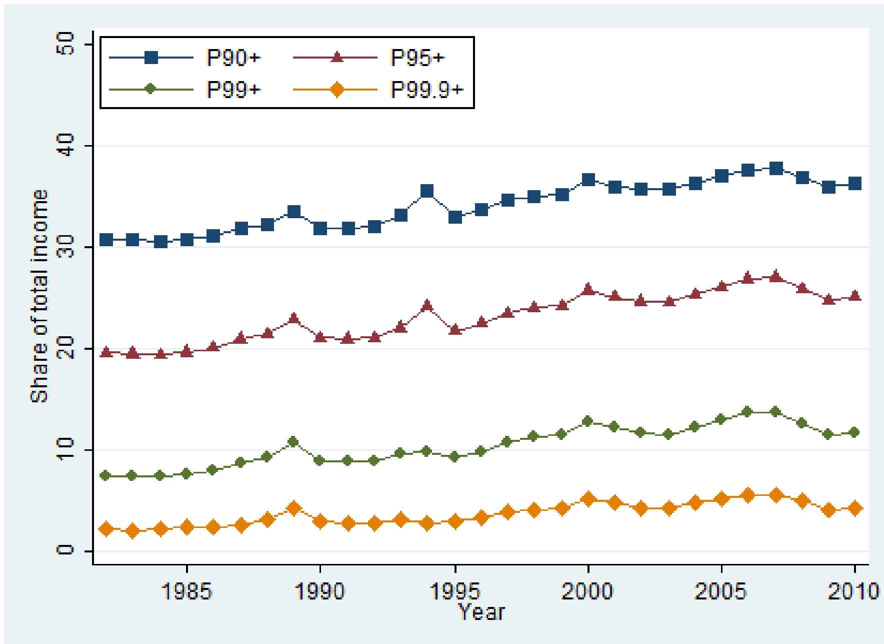

Figure 1: Top income shares, 1982-2010

Source: CANSIM 204-0001. Shown is the share of income above the indicated cutoff percentile.

The source of our income data is the CANSIM high incomes database.[9] We take the series for total income (including capital gains) as our main data for analysis.[10] We adjust all incomes to 2010 dollar values using the Consumer Price Index. The data are aggregated to the national and provincial levels. We can observe for each of several fractiles the threshold cutoff for the fractile, the average and median income within the fractile, and the share of total income. At different places in our analysis we make use of data for those above cutoffs at the 50th, 90th, 95th, 99th, 99.9th, and 99.99th percentile of income.

To provide a sense of overall trends in high incomes in Canada, Figure 1 shows the national time series for several top income shares over the period 1982

[9] CANSIM table 204-0001 provides high income threshold cutoffs and income totals for several measures of income covering the years 1982 to 2010. The ultimate source of these data are the Longitudinal Administrative Databank, a ten-percent sample of Canadians drawn from tax records.

[10] The database also includes 'market' income which excludes transfer income. Both total and market income are available with and without capital gains included. Our results are little changed when one of these alternatives is used in place of total income with capital gains.

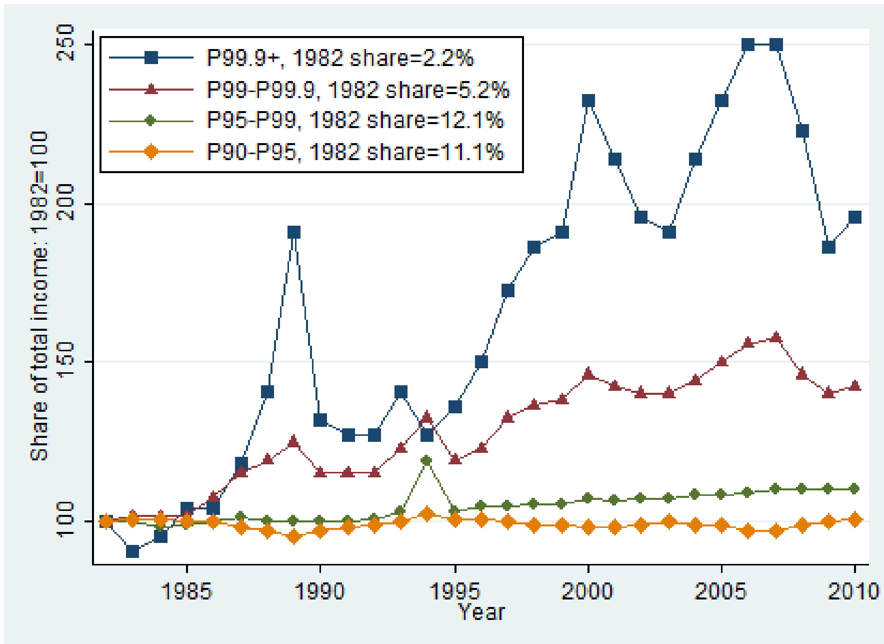

Figure 2: Top income shares, 1982-2010

Source: CANSIM 204-0001. Shown is the share of income between the indicated cutoff percentiles.

to 2010. There is clear growth above the 90th, 95th, 99th, and 99.9th percentile. However, the overall share of income going to the top half of the distribution (not shown here) stayed remarkably stable, moving only from 82.2 percent in 1982 to 83.3 percent by 2010. Given that the top ten percent grew from 30.6 to 36.2 percent over that same time period, this suggests a reshuffling of income shares within the top half of the income distribution toward the top. Moreover, the percentage growth in the fractiles shown in Figure 1 was higher for the top fractiles. In the top ten percent (above the 90th percentile cutoff; P90) the share of income grew 18 percent between 1982 and 2010, from 30.6 percent of total income to 36.2 percent. Above P95, the growth was stronger at 28 percent over this time period; for the top 0.1 percent (above P99.9), the total income share almost doubles from 2.2 to 4.3 by 2010.

The strength of the income share movements at the top of the distribution become more stark by further parsing the data. Figure 2 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 slight growth from 12.1 percent in 1982 to 13.3 in 2010. It is only in the top one percent where substantial growth is seen. For the first 9 tenths of the top one percent, the income share grows from 5.2 to 7.4 percent, or 42 percent. For the top one tenth of the top one percent, the income share nearly doubles.

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. 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.

The tax rates for our analysis come from the Canadian Tax and Credit Simulator (CTaCS; see Milligan 2013), which provides a calculation of income tax liability given a province, year, and a vector of income and family structure inputs. The CTaCS calculator is available for the years 1962 to 2012, which spans the years 1982 to 2010 available for the income data. Because our focus for this paper is top incomes, Canada's vast system of targeted refundable tax credits does not affect our estimates, as individuals at income levels observed in the top fractiles are out of the eligible range for these credits. We are interested in marginal tax rates rather than tax liabilities. To calculate the marginal tax rates we perform each simulation twice—once with the actual income and then again with earned income incremented by $100. We take the difference in tax liability between these two runs and divide by 100 to obtain the marginal tax rate.

With the tax calculator we also impute amounts for certain deductions and credit categories which can reduce taxable income and tax paid.[11] These imputations may be particularly important for higher earners who may have more complicated tax returns and greater opportunities to use available channels to reduce tax liabilities. The imputations matter to the extent the simulated individual is pushed into a lower tax bracket. However, the tax brackets in Canada

[11] The imputations are based on data from the Canada Revenue Agency's published 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 CRA data. We do this for 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.

over this period do not exhibit progressivity in higher income ranges. In fact, the income level required to be in the top one percent of income earners places the individual in the top tax bracket (before imputed deductions) for every province and year we consider.

Observed tax rates at the provincial level vary both because incomes differ across provinces and because statutory tax rates differ. Because we wish to regress incomes at the provincial level on these tax rates, we want to include only the statutory tax rate variation. To do so, we perform our simulations using a 'synthetic' income distribution, rather than the observed incomes for a given province and year. Specifically, we use incomes from the national distribution for the year 2000, and then adjust these incomes according to CPI to create data for each year. This common set of incomes is then put through CTaCS for each province and year combination. The set of incomes we use is the median total income within the given fractile.

Kopczuk and Slemrod (2002) caution against thinking of taxable income elasticities as invariant parameters. In particular, we should expect the elasticity to depend on the definition of taxable income—which changes when income definitions or exemptions and deductions are changed. In Canada, a major tax reform in 1988 which changed the federal personal income tax base substantially. Since 1988, the federal tax base has been quite stable. Nine of the ten provinces allow the federal government to collect taxes on their behalf through a tax collection agreement. A core element of these tax collection agreements is the use of the federal tax base for calculating provincial tax liabilities. The one exception is Quebec, although the tax base differences in practice are modest. Taking this into account, our analysis focuses on the years of stability for the tax base, spanning 1988 to 2010.

We graph the high income tax rates by province and year in Figure 3. The bottom line in the graph shows the high tax rate for Alberta. The tax rate in Alberta hovered around 45 percent through most of the 1990s, shifting both because of some changes in federal surtaxes at the beginning of the 90s and also changes in the Alberta tax rate. Heading out of the 1990s, Alberta first dropped its overall tax rate slightly in 1998, and then removed a surtax in 2000.[12] The sharp drop in 2001 to 39 percent occurred as the province shifted to a flat 10 percent tax rate. When combined with the federal top rate of 29 percent, this led to a top rate of 39 percent. This tax rate has remained unchanged through the subsequent years. The top tax rates in the other provinces are less distinct in

[12]The tax decrease in 1998 was from 45.5 percent of basic federal tax to 44.0 percent. Given the top federal rate of 29 percent, this dropped the effective Alberta basic rate by less than half of a percentage point. The surtax removed in 2000 was 8 percent of basic Alberta tax, which works out to a drop in the effective marginal rate of just over one percentage point.

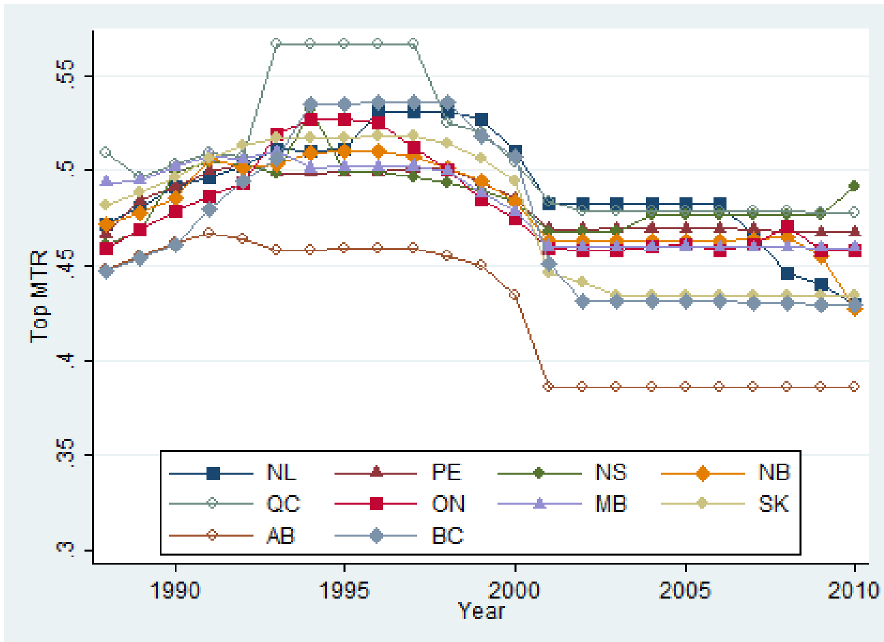

Figure 3: Top tax rates, 1988-2010

Source: Calculations using CTaCS. Shown is the top tax rate for each province through time.

Figure 3. This is a virtue for our analysis, as the cloud of lines reflect substantial changes in tax policy at the provincial level during this 23 year period. As one example, BC started with a tax slightly lower than Alberta's until 1991, then moved up to have the second highest rate at over 54 percent in the mid-1990s before falling back to the second lowest in the early 2000s at 43.7 percent. As another example, the top rate in Newfoundland and Labrador fell from the highest in the first half of the 1990s to among the lowest by 2010.

The large drop common across provinces in 1999-2001 reflects changes made at the federal level. Two separate federal income surtaxes were removed over this time period, effectively lowering the top federal tax rate by over two percent.[13] There were also substantial changes in several provinces over that time period.

This substantial variation in provincial tax rates can be summarized using a regression of the top marginal tax rate on dummies for each province and each

[13]The surtaxes were applied as a percent of the Basic Federal Tax. For high earners in the top (29 percent) federal bracket, the surtaxes of 5 and 3 percent led to a 2.3 percentage point increase in the marginal tax rate.

year in our data. The R-squared from this regression is 0.757, which suggests that the within-province through time variation accounts for around one quarter of the total variation in tax rates. This helps to justify our empirical strategy which exploits this within-province through time variation—if the within-province across time variation were small, it would be hard for our regressions to identify the impact of tax rates on reported incomes using this empirical strategy.

Our model takes special interest in the tax rate of the internal tax haven province, Alberta. In Figure 3 it appears that tax rates in Alberta have not varied substantially, outside of the large jump downward in 2001. This will present a challenge for identification in our empirical work, since it will be difficult for us to separate the impact of any other unobserved factor that varies in the pre/post 2001 period from the impact of taxes. Our main empirical specification includes several macroeconomic controls (discussed below) in an attempt to capture some determinants of income that may vary before and after 2001.

To preview the core empirical relationship we will explore in the regressions, we plot in Figure 4 the relationship between the tax rates on top income and the 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. In our estimation, we do not include observations from Alberta because of its role as the tax-haven province. 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.

Our main empirical specification when we study the impact of the internal-tax haven (Alberta) will include several macroeconomic controls. Since our specification including the Alberta tax rate will not be able to include time dummy variables, we include these macroeconomic controls to proxy for changes that may impact incomes over time. The set of controls comprises the US federal funds rate, the bank rate in Canada, the Canadian-US exchange rate, and the income share of the top one percent in the United States.[14]

5 Results

Our estimation is based on equation (5). We begin with by analyzing the taxable income responsiveness without accounting for potential shifting toward the tax-haven province. We examine the robustness of these results with different ways

[14]These variables are taken from the CANSIM data base, except the US top one percent income share, which comes from the World Top Incomes Database at the Paris School of Economics (http://topincomes.g-mond.parisschoolofeconomics.eu/).

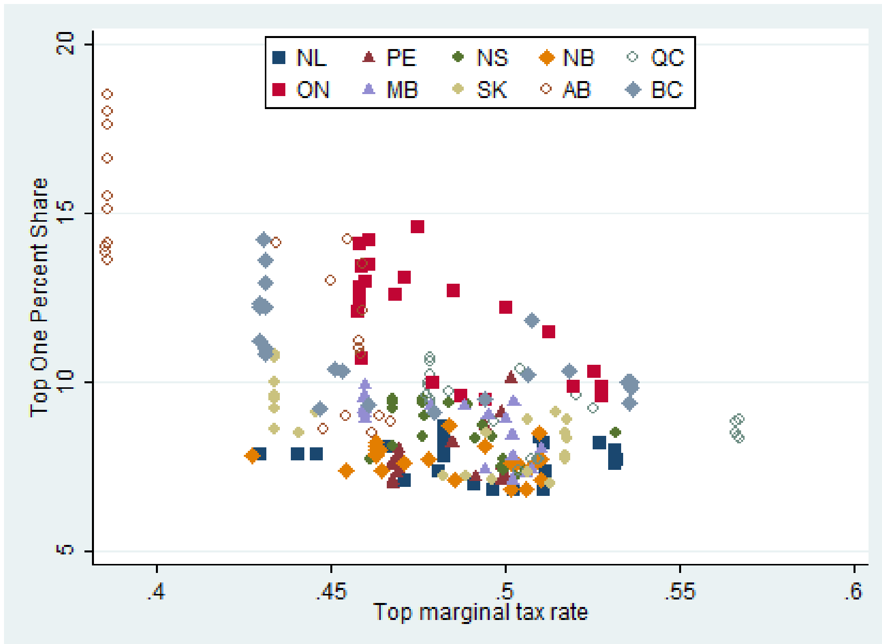

Figure 4: Top one percent tax rates and income shares, 1988 to 2010

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

of controlling for time and income trends. We then compare results across different top income shares. Finally, we move to the full specification laid out in equation (5) by including both the own-province and Alberta tax rates in the regression, and assess sensitivity to different ways of controlling for time.

5.1 Benchmark specifications for the ETI

The first set of results appears in Table 1. To facilitate comparisons with the previous literature, we look only at the own-province tax rates τ_{Hit}. So, the Alberta tax rate τ_{H0t} is left out of the estimation. We use the log of the ratio of shares as indicated in equation (5) so the resulting estimate of e_u, the coefficient on the home tax rate, can be interpreted as the semi-elasticity of taxable income. The sample includes all provinces outside Alberta for the years 1988 to 2010. The standard errors are robust-corrected for heteroskedasticity, clustering by province. We report the coefficient α, as well as the taxable income elasticity for each specification.

The first two columns in Table 1 show the results for a basic specification include provincial and year dummies. The first column shows unweighted results while the second column shows results weighted by the number of tax filers. Solon, Haider, and Wooldridge (2013) provide a discussion of when weighting is appropriate, and advise researchers to present both weighted and unweighted estimates in grouped data when attempting to show causal relationships. Since there are more residents of Ontario than of Prince Edward Island, there is a higher variance for the provincial mean in the smaller province. The lower standard error in the 2nd column when weighting indicates a gain in efficiency when using the weights.[15]

The unweighted coefficient on the marginal tax rate is -1.596, which is significant at the 1 percent level. When multiplied by the average of the tax rate variable, we can recover the elasticity of taxable income implied by this estimate. Here, that estimated elasticity is -0.774. In the second column, with the weights, drops the coefficient down to -1.180, but because of the efficiency gain the *t*-statistic is larger. The estimated elasticity falls to -0.572. This baseline estimate is remarkably similar to the elasticity of 0.74 estimated by Saez, Slemrod and Giertz (2010) for the United States using a similar methodology. Saez and Veall (2005) in turn obtained an elasticity of 0.48 for Canada, but their estimate is identified from variation in the federal tax rate alone, and so does not incorporate the potential effect of cross-province income shifting on reported incomes.[16] Atkinson and Leigh (2010), applying a methodology very similar to ours to estimate the ETI using top income shares in five Anglo-Saxon countries, obtain a semi-elasticity of 1.23, again very similar to our estimates.

In column three of Table 1 we add the log of total income. This follows a common practice in the literature (e.g. Atkinson and Leigh, 2010) to include a broad measure of income growth. Total income growth has a clear positive relationship with income shares, but has only a small impact on our estimated tax coefficient, which falls to -1.021. The accompanying taxable income elasticity lies at -0.495.

The next two columns include richer controls for time trends in the dependent variable. In column 4 we include a linear trend for each province—this is in addition to the provincial fixed effects and year fixed effects. Our point estimate is little-changed, although the standard error increases by 82 percent relative to column 3. In column 5 we introduce a quadratic trend for each province. At this

[15] In the literature some advocate weighting by income in order to get a dollar-weighted population estimate. We did not pursue that approach because of the endogeneity of observed income with the tax rates.

[16] As well, their estimates incorporate a different set of controls for non-tax factors influencing the distribution of income.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Observations	206	206	206	206	206	108	89	98
R-Squared	0.920	0.960	0.980	0.980	0.990	0.960	0.980	0.990
MTR 99th percentile	-1.596***	-1.180***	-1.021**	-1.145*	-0.883	-1.372	-0.780	-0.909*
	[0.469]	[0.290]	[0.337]	[0.613]	[0.599]	[0.762]	[1.715]	[0.421]
Elasticity	-0.774	-0.572	-0.495	-0.555	-0.428	-0.665	-0.378	-0.441
Log total income			1.095***	1.147***	1.483***	0.852*	1.148***	1.717***
			[0.259]	[0.297]	[0.327]	[0.378]	[0.114]	[0.221]
year dummies	yes	yes	yes	yes	yes	yes	yes	yes
Provincial trend	no	no	no	linear	quadratic	no	no	no
Year range	1988	1988	1988	1988	1988	1988	2001	1995
	2010	2010	2010	2010	2010	1999	2010	2005
Weights	no	yes	yes	yes	yes	yes	yes	yes

Note: In brackets are robust standard errors, clustered by province.

Table 1: Basic top income elasticity results

	P90+	P95+	P99+	P99.9+	P99-P99.9
	(1)	(2)	(3)	(4)	(5)
Observations	207	207	206	159	159
R-Squared	0.970	0.980	0.980	0.960	0.979
MTR at threshold	0.086	-0.187	-1.021**	-1.970**	-0.617**
	[0.424]	[0.334]	[0.337]	[0.651]	[0.237]
Elasticity	0.038	-0.087	-0.495	-0.962	-0.299
Log total income	0.708***	0.789***	1.095***	1.597***	0.624**
	[0.195]	[0.220]	[0.259]	[0.268]	[0.277]

Note: In brackets are robust standard errors, clustered by province.

Table 2: Results across different fractiles

point, our point estimate drops and we lose statistical significance. In further experimenting with cubic and quartic trends it is clear our specification cannot be pushed further than the linear provincial trend. However, in this type of empirical setting, we consider the result quite robust for time controls.

The final three columns experiment with different time groupings. In column 6, we look only at the period 1988-1999; in column 7 we include the grouping 2001-2010. These groupings are natural given the large shock common to many provinces in 2000 that is evident in Figure 3. We revert here to a specification like column 3 with the year and province dummies, but without province-specific trends. For the 1988 to 1999 period in column 6, the standard error and estimate grow, leaving it statistically insignificant at conventional levels (the $/italp-$value is 0.109). For the 2001-2010 period the estimate is also insignificant. This suggests that the large degree of variation surrounding the 2000-2001 years is quite important to our estimation. We narrow our sample down to the ten-year period 1995-2005 surrounding this change in column 8, finding an estimate of -0.909, which is quite similar to the -1.021 seen in column 3.

The next set of results compares the income responsiveness of different fractiles of income. We use the same specification as was used in column 3 of Table 1, including year and province dummies as well as the log of total income. The first four columns of Table 2 move from using the income share of P90 and above to using only those above P99.9 in the top tenth of one percent. The results for P90+ and P95+ indicate no responsiveness to tax rates. Column 3 of this table reproduces the results of column 3 of the previous table for those in the top 1 percent. Finally, in column 4, the estimate of -1.970 implies an elasticity of -

0.962. This suggests a very strong responsiveness in the top tenth of one percent of the income distribution. In the final column, we extract the income share of those in the bottom nine tenths of the top one percent. Here, the coefficient is lower, at -0.617, but statistically significant. The results in the last three columns line up with the intuition that the elasticity should be highest among those with the most flexibility and with access to the best advice.

5.2 Shifting and avoidance elasticities

We begin in Table 3 to explore our full specification from equation (5), which includes τ_{H0t}, the Alberta tax rate. This Alberta tax rate doesn't vary across provinces, so it is not possible to include both τ_{H0t} and the set of time dummy variables. Instead, we control using a flexible polynomial in time and the macroeconomic control variables. The first four columns in the table investigate how well the time polynomial and macroeconomic controls do in controlling for time trends, compared to the base time dummies specification. This investigation is undertaken without include the τ_{H0t} Alberta tax rate in order to calibrate our time controls as best possible before introducing τ_{H0t}.

The first column replicates the standard specification from column 3 of Table 1. In the second column, we introduce a quartic function in time, plus our set of macroeconomic controls. The point estimate is little changed, moving to -0.995. In the third and fourth column we show what happens when we include province-specific linear and quadratic trends. The result is little changed when linear trends are present in column 3, but loses statistical significance when the quadratic provincial trends are included in column 4. We conclude from this investigation that our quartic in time plus macroeconomic controls do well in replicating the results we saw with the time dummies included.

The final column of Table 3 includes the Alberta tax rate, τ_{H0t}. In this specification, the estimated coefficient on the own-province tax rate is -1.043, while the coefficient for the Alberta tax rate is 0.662. Both are statistically significant. Using the decomposition of the estimated own-rate as indicated in equation (4), we can calculate the shifting semi-elasticity and the avoidance semi-elasticity as 0.662 and 0.381, respectively.

We explore the robustness of the result including the Alberta tax rate in Table 5.2. The first column reproduces the results of Table 3, column 5. In the second column, we introduce a linear provincial trend. The results are little changed. As with earlier specifications, when we include quadratic or higher provincial trends (not shown here) the results fade to statistical insignificance. In the third column, we focus on observations only in the range 1995 to 2005. Here, the statistical significance of the Alberta tax rate coefficient is lost, but the

	(1)	(2)	(3)	(4)	(5)
Observations	206	206	206	206	206
R-Squared	0.980	0.970	0.980	0.980	0.970
MTR 99th threshold	-1.021**	-0.995**	-1.067*	-0.813	-1.043**
	[0.337]	[0.320]	[0.522]	[0.532]	[0.346]
Alberta 99th threshold					0.662***
					[0.190]
Log total income	1.095***	1.114***	1.163**	1.489***	1.119***
	[0.259]	[0.301]	[0.361]	[0.396]	[0.305]
Elasticity	-0.495	-0.482	-0.517	-0.394	
Shifting semi-elasticity					0.662
Avoidance semielasticity					0.381
time dummmies	yes	no	no	no	no
quartic in time	no	yes	yes	yes	yes
macro controls	no	yes	yes	yes	yes
prov trends	no	no	linear	quadratic	no

Note: In brackets are robust standard errors, clustered by province.

Table 3: Controlling flexibly for time, including Alberta tax rate

	(1)	(2)	(3)	(4)	(5)
Observations	206	206	98	206	159
R-Squared					
MTR at threshold	-1.043**	-1.126*	-0.913*	-1.123***	-2.238***
	[0.346]	[0.565]	[0.410]	[0.326]	[0.537]
mtrab99	0.662***	0.706**	0.818	0.705***	1.916***
	[0.190]	[0.280]	[1.094]	[0.185]	[0.342]
log total income	1.119***	1.175**	1.706***	1.127***	1.710***
	[0.305]	[0.365]	[0.226]	[0.265]	[0.437]
Shifting semi-elasticity	0.662	0.706	0.818	0.705	1.916
Avoidance semielasticity	0.381	0.420	0.095	0.418	0.321
Year range	1988	1988	1995	1988	1988
	2010	2010	2005	2010	2010
Instrument for MTR	no	no	no	yes	no
Use 99.9th percentile	no	no	no	no	yes
linear provincial trends	no	yes	no	no	no

Note: In brackets are robust standard errors, clustered by province.

Table 4: Robustness of the Alberta results

point estimate is of similar magnitude. In column 4, we use the actual province-year income to calculate the marginal tax rate, and then use the fixed-income tax rate as an instrument. This specification shows similar results. Finally, we try a specification that looks only at the top tenth of the one percent, finding coefficient estimates significantly higher than for the top one percent as a whole.

6 Policy implications

Our estimates of the shifting elasticity our large, which may appear inconsistent with the relatively high levels of redistribution in provincial personal tax systems, where top marginal rates currently range from 13 to 24 per cent. Moreover, several provinces have recently increased their top tax rates, even as the federal top rate has remained unchanged since 1988.[17] Can these facts be reconciled with our elasticity estimates in the light of our model?

One way to address these issues is to compute the marginal revenue and marginal excess burden associated with unilateral increases in provincial top marginal tax rates. In our model, tax rate increases have both "horizontal" externalities, given the effect on Alberta tax revenues through cross-border shifting, and "vertical" fiscal externalities, given that pure avoidance activities shrink the federal base for personal income tax revenues as well as the provincial base.

Recall that taxable income in province i is

$$y_i(\tau, \tau_0) = z - a(\tau_i) - s(\tau_i - \tau_0)$$

Provincial tax revenues are

$$r_i(\tau_i, \tau_0, T) = (\tau_i - T)[y(\tau_i, \tau_0) - k] + \bar{\tau}_i k \tag{6}$$

where $\tau_i - T$ is he provincial top tax rate in province i. Total tax revenues comprise provincial and federal revenues, and also revenues accruing to the low-tax province due to shifting. Accordingly, total revenues paid by the representative high-income taxpayer in province i are

$$R_i(\tau_i, \tau_0) = \tau_i[y_i(\tau_i, \tau_0) - k] + \tau_0 s(\tau_i - \tau_0) + \bar{\tau}_i k \tag{7}$$

The excess burden of a unilateral tax rate is defined as

$$b_i(\tau_i, \tau_0, T) = -u(\tau_i, \tau_0) - r_i(\tau_i, \tau_0)$$

[17] Between 2010 and 2014, the top marginal tax rate is increasing in the provinces of Nova Scotia (19.2% to 21%), New Brunswick (14.3% to 17.8%), Quebec (24% to 25.75%), Ontario (17.4% to 20.5%), and British Columbia (14.7% to 16.8%).

Province	Mean Income over $250,000	Provincial MTR	Marginal revenue per dollar of taxable income: Provincial	National
	$	%	%	%
Newfoundland	421,443	13.3	26.8	22.4
Prince Edward Island	383,538	18.4	15.6	11.2
Nova Scotia	450,065	19.2	24.4	20.0
New Brunswick	419,414	14.3	25.5	21.0
Quebec	467,970	24.0	21.5	22.1
Ontario	559,529	17.4	37.2	32.7
Manitoba	479,261	17.4	29.7	25.3
Saskatchewan	466,127	15.0	30.7	26.3
British Columbia	560,068	14.7	40.0	35.6
National Average	532,548	17.2	35.1	30.7

Table 5: Simulated effect of unilateral tax increases on revenues

from the provincial perspective, and

$$B_i(\tau_i, \tau_0) = -u(\tau_i, \tau_0) - R_i(\tau_i, \tau_0)$$

from the national perspective. Thus there is a monotone relationship between the marginal excess burden of the tax and its marginal revenue. To present results, it is convenient to scale marginal revenues from a tax increase as a percentage of taxable income y_i. Differentiating (6)–(7), and using (2)–(3), we have

$$\frac{\partial r_i / \partial \tau_i}{y_i} = \frac{y_i - k}{y_i} - (\tau_i - T)(e_a + e_s) \tag{8}$$

$$\frac{\partial R_i / \partial \tau_i}{y_i} = \frac{y_i - k}{y_i} - \tau_i e_a - (\tau_i - \tau_0) e_s \tag{9}$$

In these expressions, the first term is the percentage by which the mean income above the bracket threshold exceeds the threshold k. As noted by Saez (2004), if top incomes follow a Pareto distribution, then this ratio is independent of k (and equal to the reciprocal of the Pareto parameter). The remaining terms represent the marginal excess burden of the tax, viewed from the subnational and national perspectives, respectively.

Simulations of the marginal revenue statistics for each province are presented in Table 6, based on income and tax rate data from 2010, and our estimated semi-elasticities. For concreteness, we simulate an increase in the provincial tax rate on incomes over $250,000; if individual taxable incomes are approximately Pareto distributed over this level, then the reported marginal revenue statistics would be approximately the same for any higher threshold. The first column of the table reports the mean income of taxpayers with taxable incomes in excess of $250,000; it varies considerably among provinces, ranging from about $384,000 in Prince Island to $560,000 in British Columbia. The second column reports the provincial top rates in 2010, then ranging form 13.3 to 24.0 per cent. The remaining columns report the marginal revenue ratios from the provincial and national perspectives. Note that the reported marginal revenue ratio of, say, 26.8 and 22.4 for Newfoundland indicate that a one percentage point tax increase on top incomes there would increase would increase provincial and national revenues by 0.268 and 0.224 per cent respectively of those taxpayers' incomes.

Notwithstanding our large estimated shifting elasticity, provincial marginal revenues from unilateral tax increases are large. This reflects the fact that much of the marginal excess burden of the tax is borne by the federal government, through the shrinking of the shared tax base. This is the vertical fiscal externality analyzed by Keen and Kotsogiannis (2002). Moreover, national marginal revenues from unilateral tax increases are also relatively high, because revenue losses to the federal government are largely offset by revenue gains to the low-tax province Alberta, given tax rates and our estimated elasticities. Finally, note that provincial marginal revenues are much higher than average in two provinces, Ontario and British Columbia, that have recently increased marginal tax rates on top incomes. This reflects the greater inequality in top incomes there, which increases the revenue potential of the tax there. This suggests an Oatesian argument in favour of decentralizing redistribution in the Canadian federation: the technology of redistribution differs substantially among regions, given differences in income distributions. Thus our estimates rationalize in part the pattern of provincial taxes observed in the data.

7 Concluding remarks

This paper studies a model of tax avoidance that incorporates an "internal tax haven" into which individuals may be able to shift some of their income for tax purposes. The model yields estimable equations which we take to data on tax rates and reported incomes from Canadian provinces. We exploit subnational variation in income tax rates to identify both a shifting and an avoidance elasticity. Using a traditional approach that considers only the own-province tax rate, we find elasticity estimates that are consistent with the previous literature.

Notably, the elasticity estimates are driven by those in the top one percent—and especially those in the top tenth of one percent. When we allow for both interprovincial shifting and for avoidance, we find evidence that both components exist, with statistically and economically relevant responses. Our preferred specification reveals a shifting response that accounts for about two-thirds of total tax avoidance.

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

Baicker, Katherine, Jeffrey Clemens, and Monica Singhal (2012) 'The rise of the states: Us fiscal decentralization in the postwar period.' *Journal of Public Economics* 96(11), 1079–1091

Chetty, Raj (2009) 'Is the taxable income elasticity sufficient to calculate deadweight loss? The implications of evasion and avoidance.' *American Economic Journal: Economic Policy* pp. 31–52

Day, Kathleen M, and Stanley L Winer (2006) 'Policy-induced internal migration: An empirical investigation of the canadian case.' *International Tax and Public Finance* 13(5), 535–564

Feldstein, Martin, and Marian Vaillant Wrobel (1998) 'Can state taxes redistribute income?' *Journal of Public Economics* 68(3), 369–396

Giertz, Seth H (2007) 'The elasticity of taxable income over the 1980s and 1990s.' *National Tax Journal*

Gordon, Roger H, and Julie Berry Cullen (2011) 'Income redistribution in a federal system of governments.' *Journal of Public Economics*

Gruber, Jon, and Emmanuel Saez (2002) 'The elasticity of taxable income: Evidence and implications.' *Journal of public Economics* 84(1), 1–32

Johannesen, Niels (2010) 'Tax evasion and swiss bank deposits.' *Manuscript, Department of Economics, University of Copenhagen*

Keen, Michael J, and Christos Kotsogiannis (2002) 'Does federalism lead to excessively high taxes?' *The American Economic Review* 92(1), 363–370

Kleven, Henrik, Camille Landais, and Emmanuel Saez (2010) ‘Taxation and international migration of superstars: Evidence from the european football market.’ Technical Report, National Bureau of Economic Research

Mintz, Jack, and Michael Smart (2004) ‘Income shifting, investment, and tax competition: theory and evidence from provincial taxation in canada.’ *Journal of public Economics* 88(6), 1149–1168

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

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

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

Young, Cristobal, and Charles Varner (2011) ‘Millionaire migration and state taxation of top incomes: Evidence from a natural experiment.’ *National Tax Journal* 64(2), 255–284