UNIVERSITÉ DE MONTRÉAL – DÉPARTEMENT DES SCIENCES ÉCONOMIQUES

The Responsiveness of Government Revenue to Marginal Tax Rates

Mémoire ECN 6007

Ian Barrett
Matricule 20000507

Sous la supervision du professeur William McCausland

Février 2016

Ce document utilise des données fiscales et démographiques pour calculer les changements dans les recettes du gouvernement engendrés par les ajustements dans le taux marginal d'imposition, et cela en mettant l'accent sur la fourchette d'imposition la plus élevée. La portée de l'étude est une sélection de pays de l'O.C.D.E. Une analyse des changement de comportement des contribuables et des différentes alternatives dont le gouvernement dispose en termes de politique fiscale en suivaient. En fin, les possibles faiblesses dans des techniques de référence sont examinées en détail.

Mot clés : Imposition, marginal, recettes, budget, revenu

This thesis uses demographic and income tax data to calculate changes in government revenue following adjustments to marginal tax rates, with a focus on the highest tax bracket. The study's scope is a variety of O.E.C.D. countries. An analysis of behavioral responses of tax payers and the options available to government in terms of taxation policy follows. Past measurement techniques are also scrutinized.

Key words : Taxation, marginal, accounts, budget, revenue

CONTENTS

INDEX OF TABLES

INDEX OF FIGURES

List of Abbreviations

A.G.I.	Adjusted Gross Income
C.A.N.S.I.M.	Canadian Socio Economic Information Management System
C.A.S.E.N.	Encuesta de Caracterización Nacional (National Characteristics Survey)
C.D.F.	Cumulative Distribution Function
C.G.R.	Capital Gains Realized
E.M.T.R.	Effective Marginal Tax Rate
E.T.I.	Elasticity of Taxable Income
E.U.	European Union
G.D.P.	Gross Domestic Product
M.T.R.	Marginal Tax Rate
N.W.	Newey-West
N.Z.	New Zealand
O.E.C.D.	Organization for Economic Cooperation and Development
S.U.R.E.	Seemingly Unrelated Regression
W.T.I.D.	World Top Incomes Database

REMERCIEMENTS

I would like to first thank my supervisor Prof. William McCausland for his help and guidance throughout this project. He has been a wonderful advisor, with very constructive comments.

Joshua Lewis, Norman Ferns and Feras Belabbas have made extremely useful contributions, either via discussions, helping me track down data, or both.

Prof. Fred Campano's assistance with income distributions and computerized integration methods was invaluable.

Lastly and most importantly, my most heartfelt appreciation to my wife Silvia for her support and patience during all of my post-graduate studies.

1. INTRODUCTION

A central question in the implementation of government programs is financing.

A substantial part of revenue raised by most national and subnational governments in the Organization for Economic Cooperation and Development (O.E.C.D.) is through income tax. Hence taxes are not only a part of fiscal policy via their direct effects on the economy. They also underline all spending considerations. Governments must decide not only which tax rates to set, but also how many deductions to allow, and how to balance personal, corporate, payroll and sales taxes.

In terms of personal income taxes, one of the most important considerations is marginal tax rates – the percentage that the government will keep of the next dollar earned by an individual. A simplistic model for calculating changes in government revenue is to multiply the average individual's amount of income subject to the tax increase by the number of individual tax payers, and to multiply the result by the percentage point change in the marginal rate being implemented.

Yet such a calculation would almost certainly be erroneous because it assumes away behavioural responses. Individuals declare income subject to taxation and have a variety of possibilities to adjust such amounts - via working less, taking greater advantage of deductions and exemptions, controlling the timing of compensation, or engaging in illegal tax evasion. Retirement and residency decisions may also be adjusted due to tax considerations. Hence income declared will likely fall when taxes are raised, and rise when tax rates are cut.

How revenue forecasting can take such behavioural factors into account is essential in the elaboration of accurate budgets, yet doing so effectively is extremely difficult.

Although a large body of work exists on the case of the United States, the rest of the world has been examined in far less detail. This thesis seeks to fill part of that gap, by looking at prospects for raising government revenue in Canada, Denmark, France, Germany, New Zealand, Norway, Spain, Sweden, Switzerland and the United Kingdom. Alternative techniques are used to make inferences about Chile and Finland, where the detail of data is more limited.

This thesis proceeds as follows: Section 2 reviews previous literature on government revenue estimation and the responsiveness of declared taxable income to changes in marginal tax rates. Section 3 combines measures of such responsiveness with other country-specific data to estimate how government revenue can be expected to change following adjustments to top tax rates. Section 4 experiments with parametric approximations as substitutes for empirical data on income distributions for countries where these are lacking.

Finally, popular methods for estimating behavioural responses to changes in taxation rates are scrutinized. Section 5 examines the bias in estimates of taxable income responses to changes in tax rates introduced by the omission of capital gains from measures of taxable income, a practice widespread in the literature. Section 6 investigates the bias in estimates of such responses caused by not accounting for migration responses, another simplifying technique widely used. Section 7 concludes.

My findings are that parametric estimation methods of income distributions often work quite well and that the omission of capital gains from measures of taxable income may introduce substantial bias to published results. However, I find no evidence of substantial migration responses to changes in top marginal tax rates.

A recurring theme which is highlighted throughout this thesis is that governments are able to significantly influence how responsive taxable income is to changes in tax rates by way of availability of deductions and ease of tax avoidance. This in turn heavily impacts the amount of revenue collected.

2. LITERATURE REVIEW

A) An Equation for Estimating Changes in Government Revenue via Changes in Marginal Tax Rates

Saez (2004) proposes a decomposition which predicts changes in government revenue in a single year following a change to the top marginal tax rate. Thus only income which is taxed at the top marginal rate is examined. Whether changes to income at the extensive margin (i.e. retirement decisions and migration decisions) are incorporated is discussed shortly.

The equation is:

$$\Delta revenue = N \bullet \Delta EMTR \bullet (z - \dot{z}) \bullet \{1 - ETI \bullet [z/(z - \dot{z})] \bullet [EMTR/(1 - EMTR)]\} \quad (1)$$

Here, N is the number of taxpayers in the top marginal rate bracket (which is held constant as a simplifying assumption) and $EMTR$ is the effective marginal tax rate, or the portion of an extra dollar of gross income which will be paid to the government. z is the average taxable income for those in the top rate bracket, and $\dot{z}$ is the level of taxable income where the top tax rate kicks in. Finally, ETI is the elasticity of taxable income for those in the top tax bracket.

The E.T.I. is defined as the percentage change in reported taxable income for the average individual in the top income tax bracket when the marginal personal net-of-tax rate increases by 1%. The net-of-tax rate is in turn defined as 1 - (effective marginal tax rate), or as the amount of the next dollar earned which the taxpayer will be able to retain.

Intuitively, the E.T.I. is simply a measure of the response of declared taxable income to changes in the portion of additional gross income that an individual keeps.

It is noteworthy that equation (1) can be decomposed into a mechanical component (the term just to the left of the main parentheses) and a behavioural term (the part subtracted from 1 within the main parentheses). The mechanical response is what the change to government revenue would be if the E.T.I. were zero, which is to say that taxpayers would not adjust their taxable income when marginal tax rates change. The behavioural response incorporates adjustments of taxable income into revenue forecasting.

Looking more closely at the behavioural response

$$\{ETI \bullet [z/(z - \dot{z})] \bullet [EMTR/(1 - EMTR)]\} \quad (2)$$

we see that this term effectively skews the true measure of change in revenue from that of the strict mechanical response. This can either lower increases in revenue or turn them into losses depending on the magnitude of the E.T.I. term.

Examining the effects of the country-specific variables in detail, it is clear that increases in average taxable income and decreases in the level of income where the top tax rate kicks in will raise the mechanical change in revenue collected.

In terms of behavioural responses, using algebraic analysis we can see that all else held constant, increases in the E.T.I. will increase behavioural responses. This is also true for the E.M.T.R. and the income threshold for where the top marginal tax rate kicks in. The opposite is true of the level of average taxable income of those in the top tax bracket.

Hence we see that raising the average taxable income of those in the top tax bracket will unambiguously cause government revenues to increase at a faster rate. The same is true for policies that cause the E.T.I. to fall. As the E.M.T.R. increases, the rate of increase in government revenue will fall, with government revenue falling if equation (2) exceeds zero.

Finally, the situation becomes ambiguous for changes to the income level where the top marginal tax rate takes effect. Decreases in this variable cause both the mechanical and behavioural terms in equation (1) to increase, and so the net effect on the change in revenue collected following a decrease in $\dot{z}$ can only be determined by taking into account the values of all the variables in the equation.

Revenue is maximized when equation (2) equals 1, which is to say that raising the marginal tax rate further will cause revenue to fall.

Such a revenue-maximizing E.M.T.R. is referred to as the Laffer rate. A Laffer response is defined as tax hikes causing government revenue to fall due to the E.M.T.R. being above the Laffer rate. A Laffer curve is a graphical representation of the relationship between E.M.T.R.s and total government revenue.

Giertz (2009) uses equation (1) to estimate revenue-raising prospects for the U.S. Table 1 reproduces his results:

Table 1

Consequences of Various Values of the Elasticity of Taxable Income on American Taxation Policy
Top Marginal Taxation Rate Bracket – 2005
(As Calculated by Giertz (2009))

Elasticity of Taxable Income	Percentage of Potential Revenue Gains Lost to Behavioural Responses	Laffer Rate
0	0.00	1.000
0.2	24.55	0.775
0.4	49.13	0.634
0.6	73.56	0.536
0.8	98.06	0.464
1	122.53	0.410

E.M.T.R. for Top Income Bracket in 2005: 0.407
Income cut-off for top marginal rate: $326,451
Average income of those in the top marginal tax bracket: $740,481

An increase in the E.T.I. of just 0.2 is associated with an additional 12 to 24 percentage points of expected revenue gains lost to behavioral responses. The Laffer rate falls very quickly as the E.T.I. rises, by as much as 22.5 points for a 0.2 increase in the E.T.I.

B) The Challenges of Estimating the Elasticity of Taxable Income

Returning to the characteristics of the E.T.I., there are several important points which should be considered.

Generally, E.T.I.s are estimated for the short to medium term, as identification is extremely difficult in the long term due to a plethora of factors which could confound the results. These range from recessions to changes in the economic outlook for specific industries, as well as time-varying impacts of control variables in the regression equation for the E.T.I. (Saez, Slemrod & Giertz (2012)). This however imposes limitations on the accuracy of estimates for the present value of future revenue when looking at the long term. We therefore need to recognize that present abilities to forecast revenue are restricted to the medium term at best.

Lindsey (1987) provided a modern introduction to E.T.I.s, with Feldstein (1995) popularizing the topic further. Their estimates were very high compared to more contemporary findings, in the range of 1 to 3 for the general population, implying that the U.S. was on the wrong side of the Laffer curve.

Lindsey's approach was to use repeated cross sections of taxpayers to predict what income distributions would have been like in 1982 had the tax schedule remained the same as in 1979. He then associated any deviations from the predicted income distribution as being due to behavioural responses to the tax reforms of 1981.

Feldstein was the first to estimate the E.T.I. using traditional panel data. This involved a collection of 4000 non-stratified, randomly selected tax returns he obtained from the I.R.S. for the years around the tax reduction of 1986. His high estimates were largely driven by the reactions of the 57 taxpayers in his sample who were subject to the highest marginal rates of 1985, and whose taxable income varied significantly over the period studied. Whether reliable inferences about large groups of taxpayers can be made from the behavioural responses of 57 individuals is not clear.

In addition to the small number of high-income taxpayers in the sample used by Feldstein, his estimation method was called into question for several reasons that also concerned the work of Lindsey (1987):

First is mean reversion, which applies at the individual level. Those with very high incomes are much more likely to experience significant reductions in earnings in the future which are unrelated to taxation. Life cycle issues (average incomes peak at a certain age before declining) also fit into this category. Without proper controls this would bias the estimates of E.T.I.s downward for tax reductions, and upward for tax increases (Auten & Carroll (1999)).

Second, those at the top of the income distribution have on average been experiencing faster increases in their income than have the rest of the population, both in nominal terms and in percentages. This is almost certainly due to factors other than income tax. Without controlling for such factors, E.T.I. estimates will be biased upward for tax cuts and downward for tax hikes.

Although the sign of the bias due to high income trends is the opposite to that of mean reversion, there is no reason to believe that the two factors will cancel (Saez (2004)).

Endogeneity of the net-of-tax rate also arises, a case of simultaneous equations. Due to the federal tax system's progressivity, marginal tax rates increase with income. Assuming that the true value of the E.T.I. is not zero, we therefore have correlation of the explanatory variable with the error term. In order to isolate the impact of taxes on taxable income, tax rates should be imputed based on an instrumented measure of taxable income. Auten & Carroll (1999) proposed as an instrument the predicted change in net-of-tax rates assuming that inflation-adjusted income remains the same as in the base year.

Higher income groups often have more opportunities to control their taxable income. This is through choosing methods of compensation (personal income vs. corporate income, salary vs. equity), or choosing the timing of compensation (delaying being paid until after a reduction in marginal rates, and advancing their pay in anticipation of an increase in marginal rates). Hence it has been commonly proposed that the E.T.I. varies by income bracket, with the highest values corresponding to the highest earners (Gruber & Saez (2002), among many others).

The solution of Auten & Carroll (1999) for mean reversion and income share trends was to include a control for base-year income, using differencing to eliminate bias from time-constant omitted variables.

Their results were drastically lower than Lindsey (1987) and Feltstein (1995), with an E.T.I. estimate of 0.55 for the population as a whole.

Feldstein (1995) and Auten & Carroll (1999) excluded capital gains from their definition of taxable income, a technique which has been repeated in most studies on the matter since. However, it is not clear whether the assigned marginal tax rates used are those that would apply to the level of income with or without capital gains, with Feldstein (1995) and Auten & Carroll (1999) not addressing this point explicitly. Auten & Carroll in particular argued that those with high proportions of capital gains in their income are required to file using Alternative Minimum Tax calculations. Since data on such individuals are not generally available in panel form, they posited that such an omission from taxable income should not have a significant impact on results obtained from commonly-used panel data sets.

We should also attempt to control for the personal tax base, which is defined as the portion of national income which is subject to personal taxation. This is embodied by the availability of deductions on tax returns. More deductions result in a narrower tax base.

Changes to marginal tax rates are often accompanied by changes to said tax base. Tax cuts may be justified by an accompanying elimination of deductions in an effort to make the reduction in rates revenue neutral. Tax hikes may be accompanied by extra deductions in an effort to reduce the impact of the increases on certain groups. Thus isolating the effects of marginal tax rates on declared income is challenging (Slemrod & Kopczuk (2002)).

Kopczuk (2005) argues that with a very broad tax base the E.T.I. for even the highest income bracket is only about 0.17, roughly a third of previously accepted estimates. He does this by including the fraction of income subject to taxation as a control variable for tax bases. This micro-level variable is thus far the strongest attempt at finding a proxy to control for the macro-level tax base.

The results of Kopczuk (2005) show that the E.T.I. is influenced by policy. Hence governments can reduce its value by eliminating deductions and thereby broadening the tax base.

Turning to illegal tax evasion, Chetty (2009b) develops a utility maximization model for agents incorporating engagement in illegal sheltering activities. He concludes that governments have a large amount of control over the impacts of tax evasion. His model shows that the marginal revenue lost to increased tax evasion can on average be exactly recuperated via optimal audits. For this to hold, a risk neutrality assumption for taxpayers is required. The point stressed is that estimates of revenue lost to increases in marginal tax rates do not include revenue later recuperated in audit.

Chetty (2009c) demonstrates that taxpayers' reactions to large rate changes are relatively much larger than to small rate changes. The implication is that elasticity estimates based on small rate changes will be much lower than those based on larger changes.

Classical theory predicts bunching at the kink points in the taxation structure, defined as the limit before a new tax rate kicks in. We should therefore observe evidence of behavioural responses via bunching. Saez (2010) finds evidence of this mainly in terms of cut-off rates for the earned employment refundable tax credit, a reaction concentrated among the self-employed. There is also a certain amount of bunching at the level of income where taxation starts. Quite surprisingly, he finds no evidence of bunching at any other kink points. Bastani & Selin (2011) find similar results for Sweden.

Further to these developments, Le Maire & Schjerning (2012) use Danish data to show that the self-employed retain earnings in their companies, keeping their taxable income just below the cut-offs to avoid entering another income tax bracket. In fact, of an overall estimated E.T.I. of 0.5 for the self-employed, only 0.14 to 0.2 is structural. The rest of the reaction is simply income shifting (or smoothing) from one year to another. Taxpayers who do not have self-employment income do not exhibit bunching behaviour in these results.

Through all of these results, we see that the more opportunities for tax evasion (legal or otherwise), the higher the value of the E.T.I., further reinforcing the findings of Kopczuk (2005).

C) Responsiveness of Capital Gains Realizations to Changes in Tax Rates

Of the quite limited work in this area, Auerbach (1988) uses time series data to establish that there are significant short run elasticities of reported capital gains to changes in the taxation thereof. However, estimates of permanent behavioral responses were not statistically different from zero.

Auerbach, Burman & Siegal (1997) use panel data to show that long run elasticities of capital gains realisations to net-of-tax rates for capital gains are indeed very low, with little avoidance behaviour aside from temporal shifting of the realisations themselves.

Section 5 of this thesis undertakes further research on this topic.

D) Migration and Evidence of Behavioural Responses at the Extensive Margin

Despite the large majority of E.T.I. research focusing on the U.S., there has been relatively little research on state marginal tax rates and their effects on migration between states, with even less on migration between countries. However, progress has been made in certain areas.

The earliest published work is that of Feldstein & Valiant (1994), who posit that states cannot redistribute income via state income taxes. This is since taxpayers can move between states. Their conclusion has, however, been contradicted by empirical research over the years.

Knapp, White & Clark (2001) use a nested logit model to find that higher state income tax rates *decrease* emigration rates. They explain this as income taxes being a proxy for public services. However, income brackets are not differentiated, so it could be that only the residence decisions of the very wealthy are affected by state income tax rates. Such a result could be hidden by an offsetting number of people moving to the state to take advantage of better public services as the very high income earners leave. Net in-migration would show as positive under such circumstances.

Bahl, Martinez-Vazquez, & Wallace (2002) go on to find evidence that states with higher tax rates tend to have higher salaries, which could also help explain the surprising results of Knapp et al.

Bakija & Slemrod (2004) look at the migration induced by state level estate taxes, and although a small correlation is found, it is not large enough to significantly impact revenue levels.

Conway & Houtenville (1998, 2001) show a negative causal relationship of marginal tax rates on net state migration of the elderly. However, Conway & Rork (2005) argue that the causal effect is more likely to be the other way around. States with influxes of seniors tend to adjust fiscal policy to accommodate them. This result is found through time variability in migration patterns and estate taxes, with the young as a control group. There could however be a causal effect of estate taxes on the migration rates of the very rich elderly.

Coomes & Hoyt (2008) examine the effects of state income tax rates on those living in metropolitan areas that span at least two states. Oddly, marginal net-of-state-tax rates have an impact on net migration rates only in states that decide residence based on state of employment (so-called states without reciprocity, which include such metropolitan areas as Chattanooga, Kansas City, Memphis, New York, Philadelphia, and St. Louis). They hypothesize that this is because there is a larger difference in state tax rates between states without reciprocity. The percentage increase in net migration to a 1% change in net-of-tax rates is estimated to be around 0.4% for those living in multi-state spanning metropolitan areas in states without reciprocity. They further estimate that a 1% increase in sales taxes is estimated to reduce net migration to the state by 0.35%.

Bruce, Fox & Lang (2010) find evidence of inter-state transferring of income, but this is done via trusts (a common financial instrument allowing income to be taxed in another jurisdiction) and not migration. This is seen through differences in federal adjusted gross income and calculated tax bases at the state level, since income received through a trust appears on tax returns at the federal level, but not on the return of the state of residence.

Young & Varner (2011) numerically estimate the elasticity of migration to state income tax rates by examining the introduction of a 'millionaire tax' in New Jersey. That is, the percent change in net migration of the wealthy to New Jersey following a 1% change in their net-of-state-tax rate. Their estimates are in the order of 0.1.

Looking at Canadian data, Day & Winer (2006) examine the effects of provincial tax rates on migration between provinces, and find very little evidence of a causal effect. Factors such as economic and (in the case of Quebec) political situations are by far the most important factors. As with many such studies, they do not differentiate between income brackets. So again, it could be that only the highest income individuals, a very small percentage of the population, are influenced to a significant degree by tax rates when choosing their province of residence.

Milligan & Smart (2013, 2014) find evidence of income shifting between provinces as a reaction to provincial tax rates, but as with Bruce, Fox & Lang (2010) this is mostly done through trusts, and not physically moving.

Kirchgassner & Pommerehne (1996) and Liebig, Puhani & Sousa-Poza (2006) study internal migration in Switzerland and its reaction to tax rates at the canton level, and find very little evidence of substantial sensitivity. They do however find evidence of migration by young, college educated citizens. These results are stated to be too small to significantly impact revenue collection. The study does not comment on the long term fiscal implications of canton-level migration of the young, however.

Kleven, Landais, Saez & Schultz (2014) find highly elastic responses of very skilled immigrants to Denmark to reductions in tax rates of recently arrived, highly skilled workers. Their estimates are of the order of 1.5 to 2.

The first paper to deal directly with inter-country migration caused by marginal tax rates is Sanandaji (2012), who uses information publicly available in Forbes to examine the migration behaviour of billionaires. He finds that the main circumstances of migration of the super-rich are to go from a poor country to a richer one, and to move to countries with political and cultural ties. However, he finds that most billionaires stay in their home countries, with taxes having a minor impact on their choices of residence. A caveat is that he only examines physical migration, and not the flight of capital.

Finally, Kleven, Landais & Saez (2013) study the migration of European soccer players between leagues and its correlation with net-of-tax rates. Interestingly, the vast majority of players choose to play in their own countries, with little impact from local fiscal policy. The exception is more skilled (and highly paid) players, who show very high elasticities of migration to net-of-tax rates of approximately 1, compared to 0.15 for domestic players.

3. PREDICTING GOVERNMENT REVENUE

A) Data Necessary for Estimates of Revenue Lost to Behavioural Responses

In order to apply equation (1) to estimate government revenue responses to changes in marginal tax rates, we need estimates of five different parameters.

The E.T.I. is the most challenging parameter on which to obtain data. The best source is the large volume of articles which have been written specifically to estimate E.T.I.s for a variety of countries. However, the techniques used by the authors vary considerably, with some being more rigorous than others. Most follow one of the three popular regression models outlined in the appendix.

The specifics of various estimates of the E.T.I. for a variety of countries are summarized in table 3.

After obtaining estimates for the E.T.I. for each country of interest, the other pieces of data necessary to apply equation (1) are values for ***EMTR***, $\boldsymbol{z}$ (the average income of those in the top marginal tax rate bracket), $\dot{\mathbf{z}}$ (the income cut-off to be subject to the top marginal tax rate) and $\boldsymbol{N}$ (the number of taxpayers facing the top marginal tax rate).

Data on E.M.T.R.s are available directly from table I.7 of the tax database of the O.E.C.D., as are values for cut-offs to be in the top marginal tax bracket.

Getting estimates of average incomes for those in the top bracket and the number of people subject to the top rate is trickier. The closest data available generally come from the World Top Incomes Database (W.T.I.D.). It contains values for the cut-offs to be in the top 10, 5, 1, 0.5, 0.1 and 0.01 percent of income earners for a variety of countries. It also contains average incomes for each of these brackets.

The first step to obtain estimates of $\boldsymbol{z}$ and $\boldsymbol{N}$ is finding the percentage of top income earners corresponding to the group which is subjected to the top marginal tax rate.

After choosing the percentile whose cut-off is closest to that of the top marginal tax rate from among the income percentiles available on the W.T.I.D., I construct a quadratic interpolation. The other points are the next higher and lower percentiles for which data are available from the W.T.I.D, along with their income cut-offs.

The corresponding quadratic equation gives us an estimate of the percentage of the population subject to the top tax rate. Multiplying by the population gives the estimate of $\boldsymbol{N}$.

A similar process, but with average incomes in place of income cut-offs, gives an estimate of $\boldsymbol{z}$.

An exception to this is for New Zealand, where the level of detail of data available from Statistics New Zealand permits direct calculation of $\boldsymbol{N}$ and $\boldsymbol{z}$.

The year of most current data varies by country in the W.T.I.D. Table 2 summarizes this, and shows corresponding E.M.T.R.s and other fiscal data.

Table 2

Comparison of Fiscal Data Available from the W.T.I.D. across O.E.C.D. Countries

Country	Year of Most Current Demographic Data	Highest Effective Marginal Tax Rate in That Year	Average Income in the Top Tax Bracket ($U.S.)*	Income Cut-Off to Enter the Top Tax Bracket ($U.S.)*
Canada	2010	0.464	207,441.80	105,477.06
Denmark	2010	0.561	82,323.92	53,343.07
France	2004	0.550	154,279.22	88,229.96
Germany	2008	0.475	884,765.8	313,188.66
New Zealand**	2012	0.380	79,903.34	47,614.80
Norway	2011	0.478	155,511.92	87,437.44
Spain	2012	0.520	972,069.83	443,849.92
Sweden	2013	0.567	107,306.09	67,821.9
Switzerland	2010	0.414	370,067.60	180,350.72
United Kingdom	2012	0.520	478,831.64	215,753.79

***Dollar figures are adjusted for purchasing price parity**

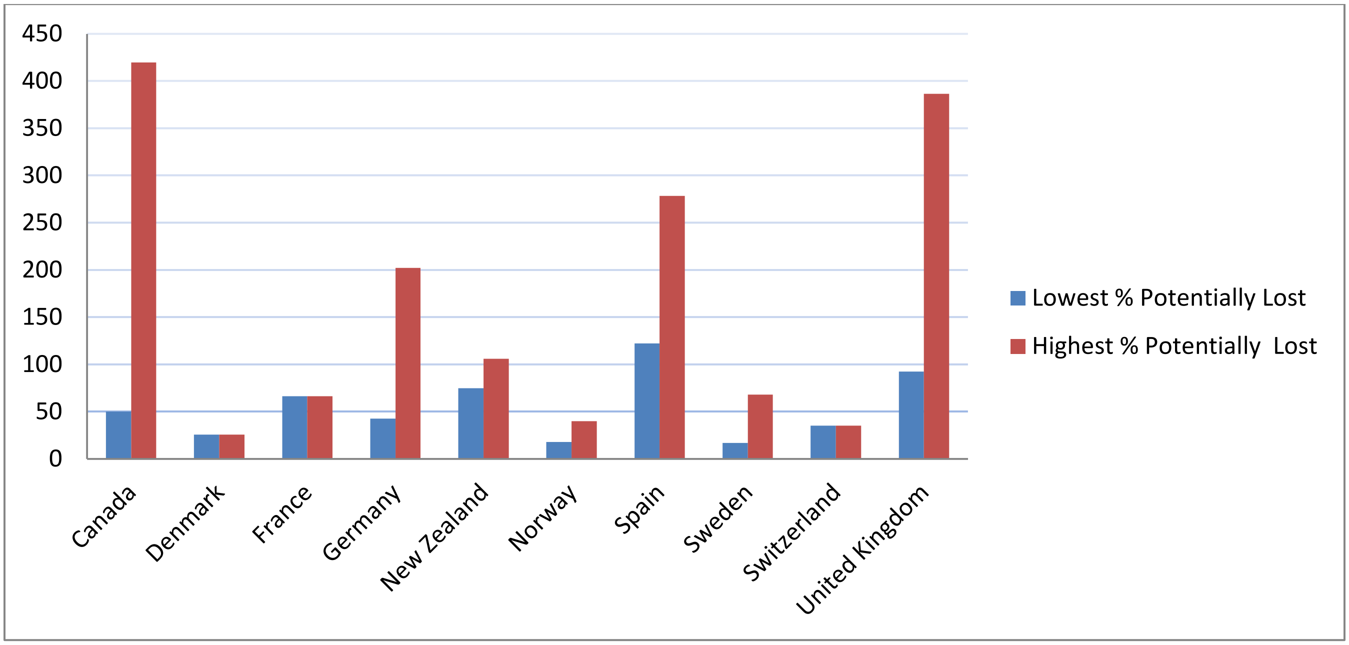
**** Data comes from Statistics New Zealand**

B) Results

The results suggest that even small differences in the estimated values of E.T.I.s can have enormous impacts on lost revenue. Combining published E.T.I. estimates with the tax situation at the time of the policy change studied, maximum and minimum values for the percentage of potential increases in revenue lost to behavioural responses following a rise in marginal tax rates are presented in figure 1:

Figure 1

Maximum and Minimum Percentages of Potential Increases in Revenue Lost to Behavioural Responses Based on Published Estimates of the Elasticity of Taxable Income



Based on these results, many countries could be on the wrong side of the Laffer curve. That is, they could raise substantial amounts of revenue by lowering tax rates. For the most extreme cases, a 1% reduction in top marginal tax rates would cause revenue collected from those in that income bracket to rise by 4%.

In table 3 on the next page, I turn to full results for potential revenue losses for each country studied, as well as corresponding Laffer rates and details of data used in the estimates.

Table 3

Potential Lost Revenue and Laffer Rates for Published Estimates of the E.T.I. by Country, Along With Pertinent Data Used

Country	Authors	Years of Study	MTR After Study	Points of Change	Model Used	E.T.I. Estimate	Standard Error	t-statistic	Sample Size	% Potential Revenue Lost	Laffer Rate
Canada	Sillamaa & Veall (2000)	1988	45.2	-5	Eq. (A2)	0.25	N/A	15.02	475,000	50.21	0.622
Canada	Gagne, Nadeau and Vaillancourt (2001)	1972 to 1996	46.69	Provincial variation	S.U.R.E.	1.08 to 3.05	0.3194 to 0.6741	3.38 to 4.53	N/A	419.81*	0.173*
Canada	Milligan & Smart (2014)	1988 to 2011	46.4	Provincial variation	Eq. (A2)	0.664	0.27	2.46	N/A	116.94	0.425
Denmark	Kleven & Schultz (2012)	1980 to 2005	63	-14	Eq. (A4)	0.04 to 0.06	0.002	20 to 30	4,000, 000	25.84*	0.868*
France	Cabannes, Houdre & Landais	1997 to 2004	47.8	-20	Eq. (A3)	0.31	N/A	N/A	500,000	66.31	0.580
Germany	Massarrat-Mashhadi & Werdt (2012)	2002 to 2004	56.3	-8	Eq. (A5)	0.16	N/A	3.20	900,000	42.55	0.752
Germany	Doerrenberg, Peichl, & Siegloch	2001 to 2008	47.5	-8	Eq. (A4)	0.364 to 0.49	N/A	9.69 to 15.67	1,000,000	68.62*	0.568*
Germany	Gottfried & Witczak (2012)	2001 to 2004	60.5	-8	Eq. (A4)	0.435 to 0.56	0.005 to 0.008	87 to 70	19,000,000	202.31*	0.431*
New Zealand	Thomas (2007)	1982 to 1990	33	-33	Eq. (A2)	0.52	0.067	7.76	829,742	66.32	0.426
New Zealand	Carey & al (2015)	1994 to 2009	38	-6	Eq. (A2)	0.5 to 0.7	0.125	4.00 to 5.60	335,190	91.00*	0.402
Norway	Aarbu & Thoresen (2001)	1991 to 1994	49.5	-10	Eq. (A2)	0 to 0.2	0.077 to 0.086	0 to 2.33	2,000	37.45**	0.724**
Norway	Thoresen (2002)	1991 to 1999	49.3	-10	Eq. (A3)	0 to 0.166	0.100	0 to 1.64	2,150	36.50**	0.727**
Norway	Vatto (2012)	1995 to 2008	47.2	6 to -7.5	Eq. (A3)	0.05	0.0021	23.81	8,000,000	39.90	0.691
Norway	Thoresen & Vatto (2013)	2000 to 2008	47.2	-7	Eq. (A3)	0.05 to 0.09	0.0039 to 0.0051	12.82 to 17.64	5,000,000	17.95*	0.833*
Spain	Diaz-Mendoza (2004)	1987 to 1994	56	****	Eq. (A3)	0.35	N/A	5.71	23,500	125.78	0.503
Spain	Arrazola et al (2014)	2006 to 2007	43	-4	Eq. (A3)	1.541	0.028	55.04	300,000	122.14	0.382
Spain	Sanz-Sanz et al (2015)	2006 to 2007	43	-4	Eq. (A2)	0.676	0.03	22.53	400,000	278.44	0.213
Sweden	Selen (2002)	1988 to 1995	59.75	-23	Eq. (A3)	0.22 to 0.534	0.09 to 0.11	2.58 to 4.94	2,000	***	***
Sweden	Hansson (2004)	1989 to 1992	53.49	-20	Eq. (A2)	0.4 to 0.5	N/A	9.86 to 21.12	80,000	***	***
Sweden	Holmlund & Soderstrom (2011)	1996 to 2002	55.5	5	Eq. (A5)	0.1 to 0.3	N/A	2.22 to 2.96	100,000	68.08*	0.647*
Switzerland	Martinez (2014)	1989 to 2008	41.4	-2	Eq. (A4)	0.24 to 0.29	0.0015	160 to 193.33	1,500,000	35.18*	0.668*
United Kingdom	Brewer, Saez & Shephard (2008)	1978 to 2003	41	20 to -30	Saez (2004)	0.46	0.13	3.54	N/A	92.54	0.429

*** Using the average E.T.I. value of the range provided in the article**
**** Using the highest E.T.I. estimate in the range provided in the article**
***** Missing sufficient demographic data to be able to obtain estimates for period in question**
****** Details of tax reform listed only as 'significant cuts for married filers'**

References to equations in table 3 are found in the appendix.

E.T.I. estimates for countries with the highest tax rates (Denmark, Norway and Sweden) are consistently low, of the magnitude 0.05 to 0.5.

Still, the percentage of revenue lost to behavioural responses is substantial, at more than 25% for even a small E.T.I. of 0.05 in the case of Denmark.

Immediately noticeable is the large degree in variation in E.T.I. estimates, not just between countries but at times within a single country. Institutional factors such as availability of deductions, cultural attitudes towards government or degrees of enforcement such as audits should explain at least part of this difference. However, methodology issues such as the regression model and corresponding assumptions made as well as the quality of data almost certainly play a role. Specifics are presented in the appendix.

Laffer rates for such countries are often substantially above the E.M.T.R.s of the period studied, suggesting that governments had a window for raising further revenue, political and welfare considerations aside.

E.T.I. estimates vary substantially, not only across countries but at times across different studies for the same country. Different estimating methods are a likely explanation for this, each with its own assumptions and weaknesses. Common estimation methods are summarized in the first part of the appendix.

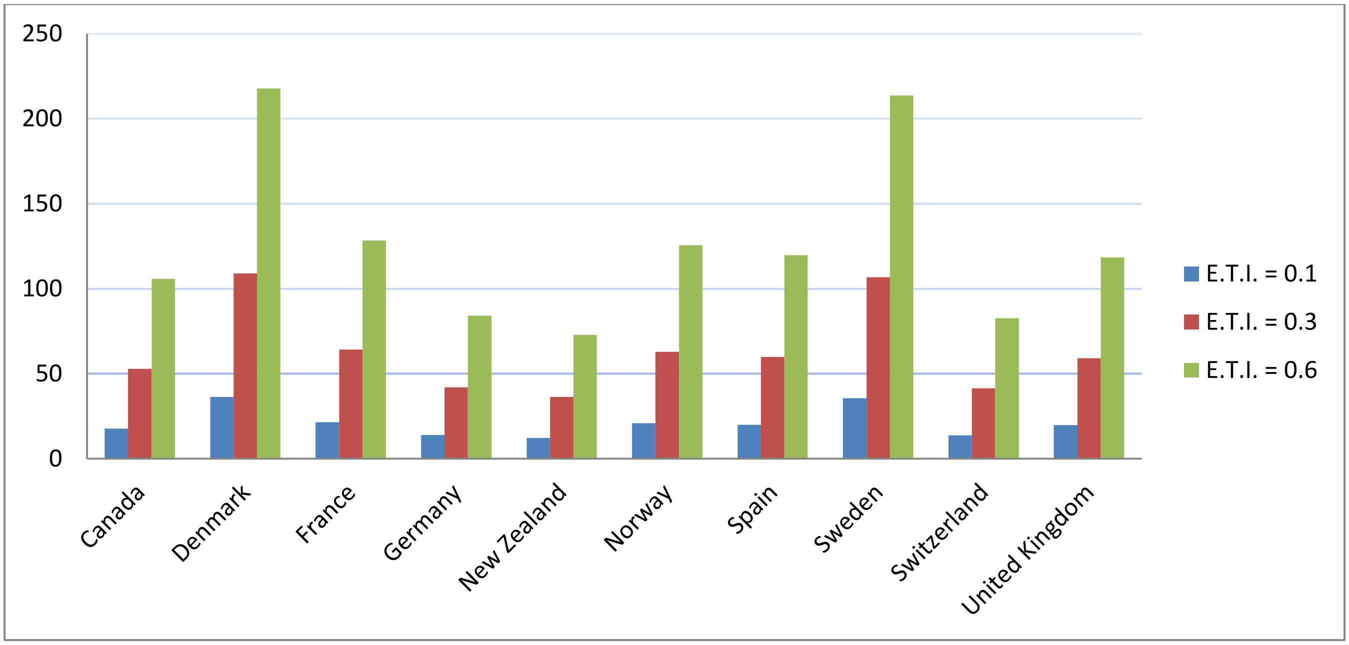
A general weakness in the data is that almost all estimates are based around tax cuts, with little done using data on tax increases. This is likely due to a global trend to lower tax rates which began in the early 1980s, the same period in which detailed data was first collected.

We see that estimates published in the last ten years have used substantially larger sample sizes, producing much more precise results. There is a general trend of high quality data becoming available across a range of European countries over the last decade.

Figure 2 shows the potential revenue increases lost to behavioural responses for a variety of countries for several plausible values of the E.T.I.:

Figure 2

Percentages of Potential Increases in Revenue Lost to Behavioural Responses for Various Values of the Elasticity of Taxable Income – Top Tax Brackets



Years of data as per table 2

The progression of the percentages lost relative to E.T.I.s is as expected given that such percentages are proportional to the E.T.I.

Figures 3 to 5 present behavioural responses and Laffer rates for Canada, Germany and Sweden.

Figure 3

Potential Revenue Lost to Behavioral Responses &
Estimates of Laffer Rates - Canada

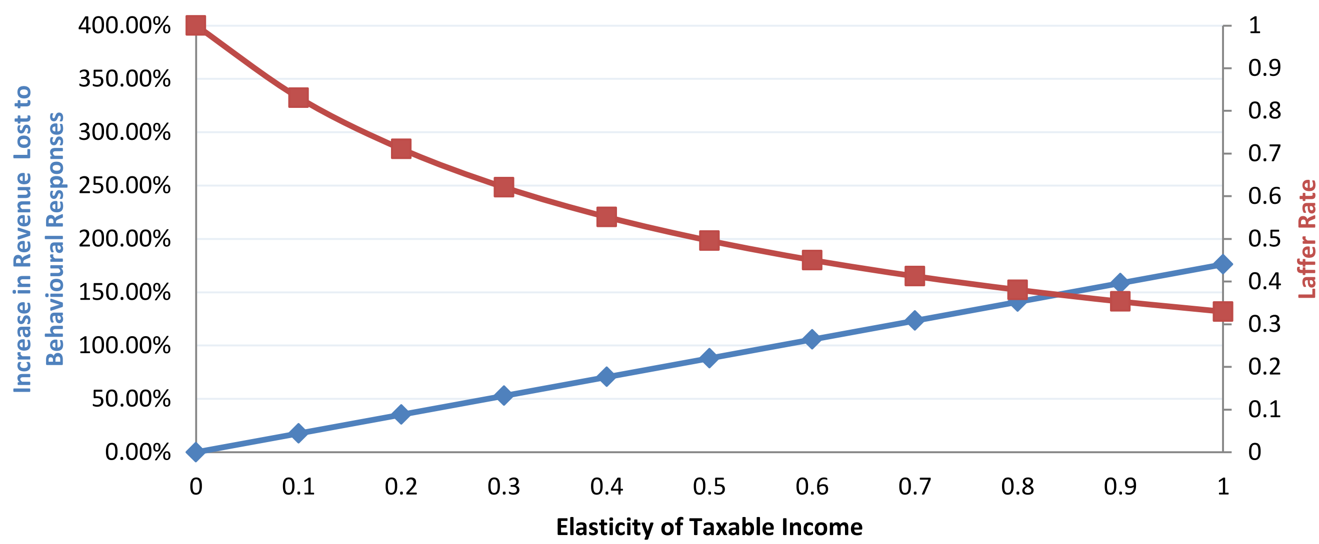


Figure 4

Potential Revenue Lost to Behavioral Responses &
Estimates of Laffer Rates - Germany

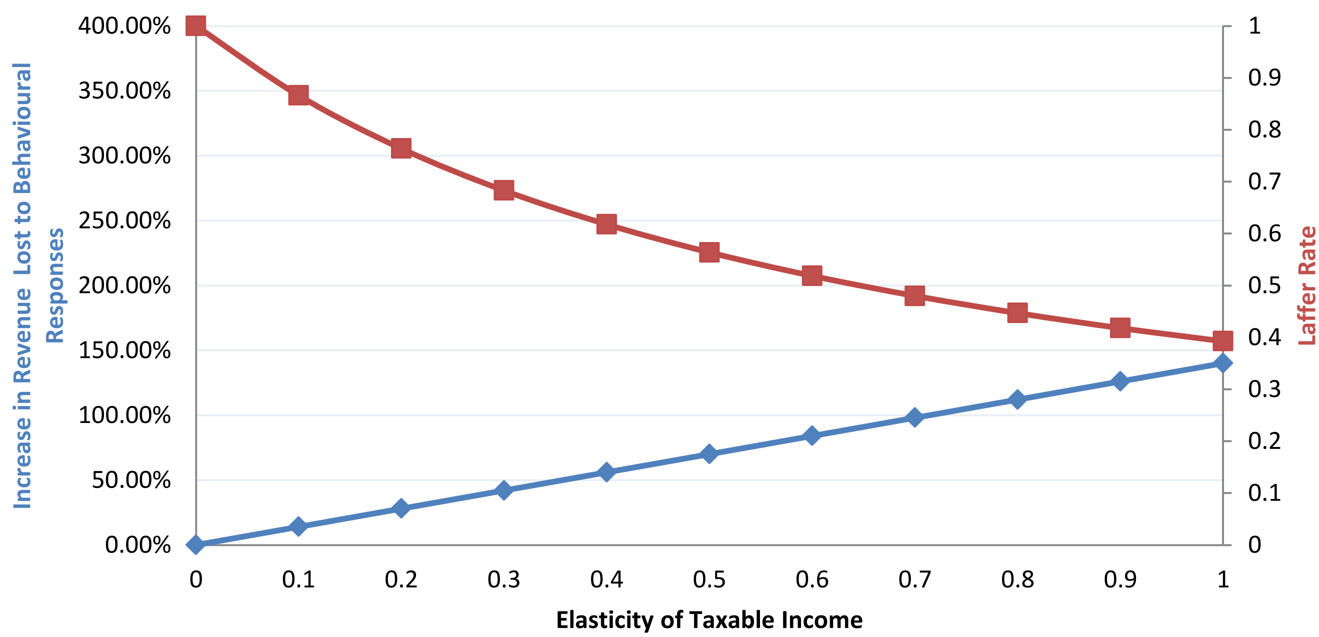
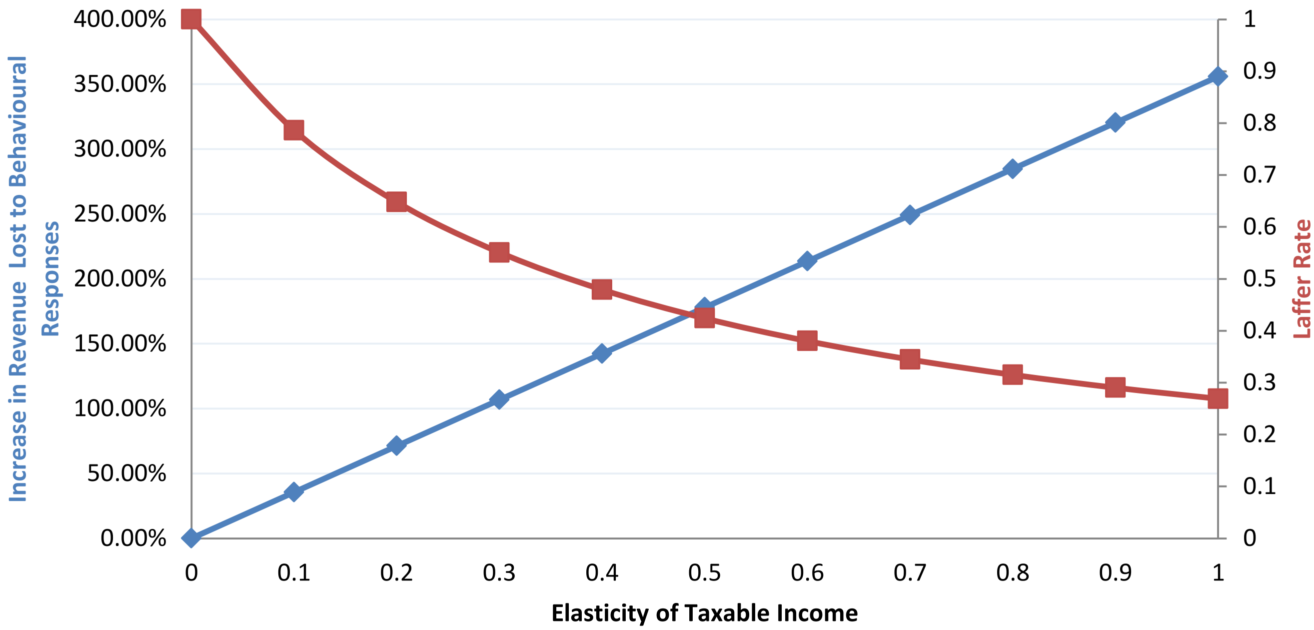


Figure 5

Potential Revenue Lost to Behavioral Responses & Estimates of Laffer Rates - Sweden



Although the percentage of mechanical revenue gains which are lost to behavioral responses varies greatly across countries, the Laffer rates are more consistent.

Further algebraic analysis of equation (1) shows that changes in N have no effect on the relationship between the E.T.I. and the percentage of increases in revenue lost to behavioral responses:

$$\Delta revenue = N \bullet \Delta EMTR \bullet (z - \dot{z}) \bullet \{1 - ETI \bullet [z/(z - \dot{z})] \bullet [EMTR/(1 - EMTR)]\}$$

We have that

i) $N \bullet \Delta EMTR \bullet (z - \dot{z})$ is the mechanical revenue response

ii) $N \bullet \Delta EMTR \bullet (z - \dot{z}) \bullet ETI \bullet [z/(z - \dot{z})] \bullet [EMTR/(1 - EMTR)]$ is the behavioral response or the amount of potential revenue lost.

iii) Therefore the percentage of revenue gains lost to behavioral responses is simply ii/i, or $ETI \bullet [z/(z - \dot{z})] \bullet [EMTR/(1 - EMTR)]$. N does not appear.

iv) Finally, $[z/(z - \dot{z})] \bullet [EMTR/(1 - EMTR)]$ is the slope of the blue lines in figures 2 to 4.

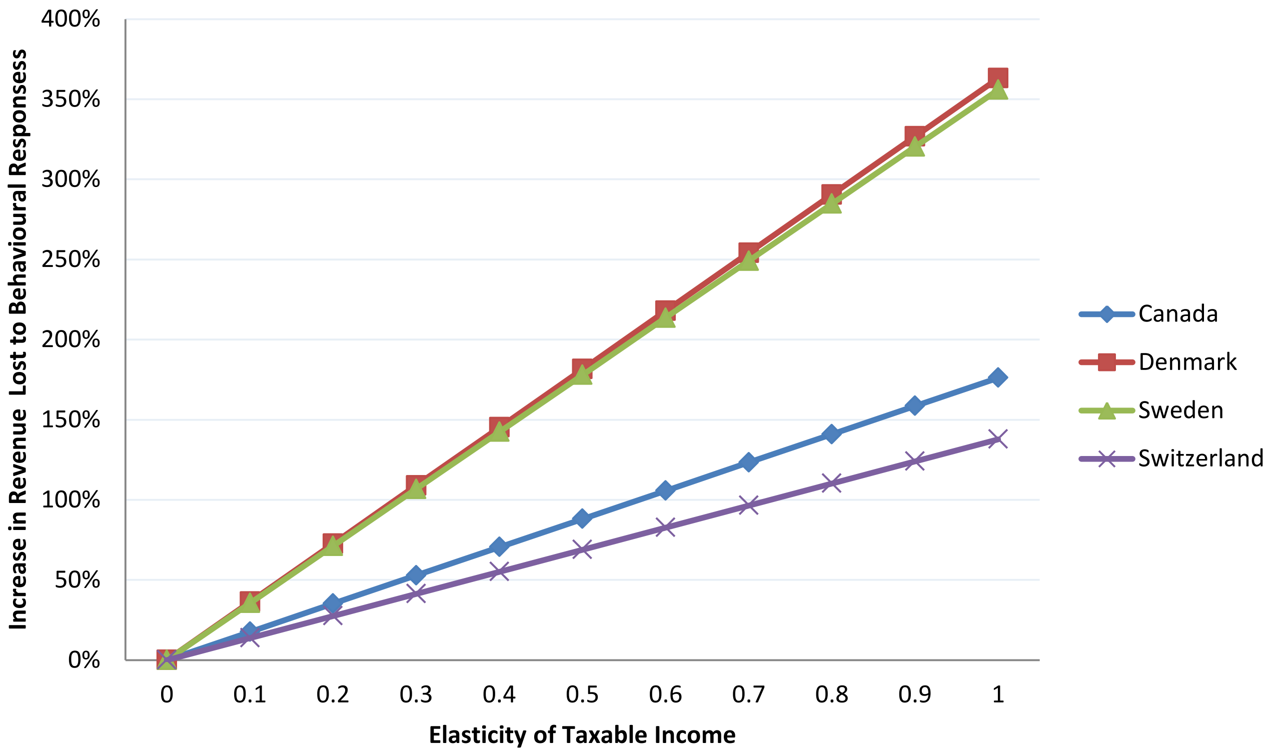
As the E.M.T.R. gets larger, the slope of the line corresponding to percentage revenue gains lost for a given value of the E.T.I. gets steeper. This is also true for decreases in z and increases in $\dot{z}$.

Hence for a given value of the E.T.I., countries that have higher percentages of potential revenue gains lost to behavioral responses either have a higher E.M.T.R., a lower average income for those in the top tax bracket, or a higher cut-off to be in the top income tax bracket. Raising average income for those in the top tax bracket will raise government revenue by increasing potential revenue to be collected. The percentage of revenue lost to behavioral responses will fall. The situation is similar when lowering the cut-off to be in the top tax bracket.

We can see these effects visually in figure 6 for a choice of 4 countries with noticeable differences in the percentages of revenue lost to behavioral responses:

Figure 6

Comparison of Potential Revenue Lost to Behavioral Reactions across Countries



Following the above the discussion, it is clear that the driving forces behind the variations across countries comes from differences in marginal tax rates, average incomes in the top tax bracket, and the level of income where the top marginal tax rate kicks in. Quantifying how much of the variation is due to each of these factors is more challenging.

A natural strategy would be to take partial derivatives of equation (2) with respect to each variable of interest. However, this strategy is confounded by the average income of those in the top tax bracket being dependent on the income cutoff to be in the top tax bracket. The higher the cut-off, the higher the average income of those left in that group. Unfortunately, the relationship between these factors is very difficult to measure without having the exact shape of the income distribution for the country in question. Making further approximations to get around this issue could lead to false inferences about which factors have the most impact.

It is perhaps preferable to simply assume that the two are independent, which should be approximately true in small neighbourhoods around the observed values.

Table 4 presents partial derivatives evaluated at the most current data points available for the four countries of figure 5, without considering the E.T.I.:

Table 4

Partial Derivatives of the Equation for Percentage of Revenue Lost to Behavioural Responses Taken With Respect to Its Influencing Factors Evaluated at Most Current Values of Each Variable

Partial Derivative W.R.T.	Canada	Denmark	Sweden	Switzerland
z	-8.80×10^{-06}	-8.10×10^{-05}	-5.70×10^{-05}	-3.50×10^{-06}
$\dot{z}$	1.73×10^{-05}	1.25×10^{-04}	9.01×10^{-05}	7.26×10^{-06}
EMTR	7.08	14.74	14.50	5.68

We see that generally, the values for all partial derivatives are higher for Denmark and Sweden than for Canada and Switzerland. The values for the partial derivatives w.r.t. z and $\dot{z}$ are easily comparable, as they are of the same units. However, it is not obvious how to compare them to the values w.r.t. the ***EMTR***.

Table 5 makes an attempt in this sense by examining ratios:

Table 5

Ratios of Partial Derivatives of the Equation for Percentage of Revenue Lost to Behavioural Responses Taken With Respect to Its Influencing Factors Evaluated at Most Current Values of Each Variable

	Canada	Denmark	Sweden	Switzerland
$\dot{z}/z$	-1.97	-1.54	-1.58	-2.05
EMTR/z	-806316	-181607	-254452	-1604616
EMTR/ ż	409984	117674	160824	782002

Here we see that in Denmark and Sweden the effective marginal tax rate and income cut-offs for the top tax bracket exert much less influence on the percentage of revenue lost to behavioural responses relative to the average income of those in the top tax bracket. Hence we can expect that the lower average incomes of those in the top tax bracket in Denmark and Sweden relative to other countries is what drives their potential for large percentages of income lost to behavioural responses.

An equation for the Laffer rate is $\mathbf{1/\{[(}\boldsymbol{z \bullet ETI}\mathbf{)/(}\boldsymbol{z}\mathbf{-}\ \boldsymbol{\dot{z}}\mathbf{)]+1\}}$. (3)

From this, we can see that decreasing the E.T.I., decreasing the cut-off to be in the highest income tax bracket, or increasing the average income of those in the top marginal tax bracket will raise the Laffer rate. Hence we expect those countries with the highest Laffer rates to have the lowest E.T.I., lowest cut-off for the top income tax bracket, or highest average income for those in that bracket. Countries with the highest tax rates such as Denmark and Sweden have the lowest Laffer rates, with Sweden's being 0.2690 for an E.T.I. of 1. This implies that the Scandinavian model of high tax rates and a wide variety of services is dependent on low E.T.I.s to function.

Figures 7 and 8 present a visual representation of Laffer rates:

Figure 7

Comparison of Laffer Rates across Countries

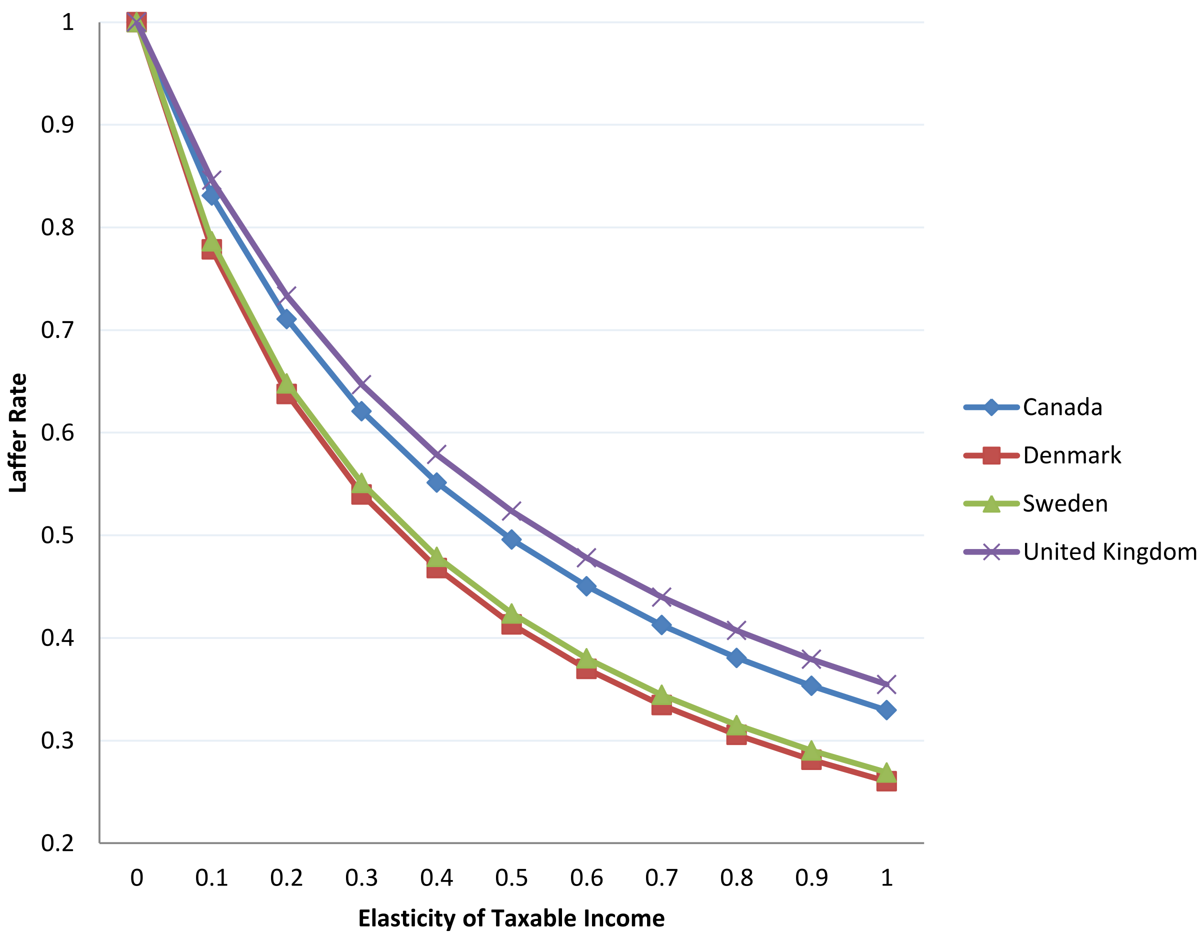
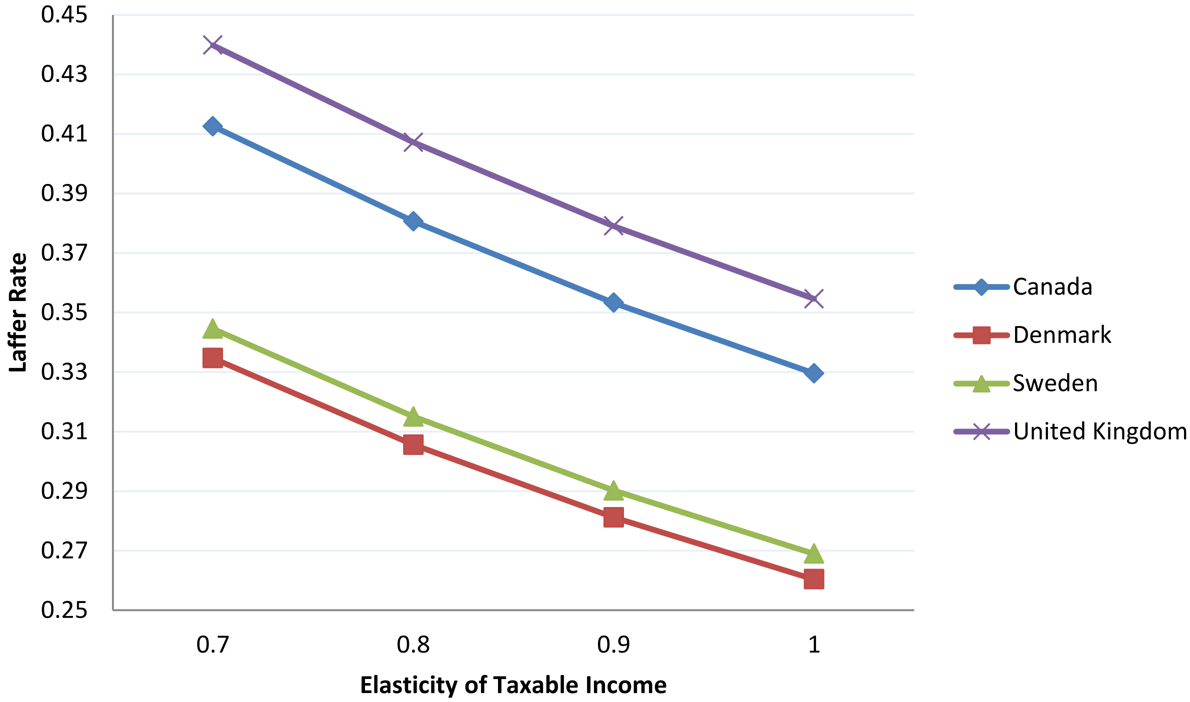


Figure 8

Comparison of Laffer Rates across Countries

Restriction to Higher Values of the E.T.I



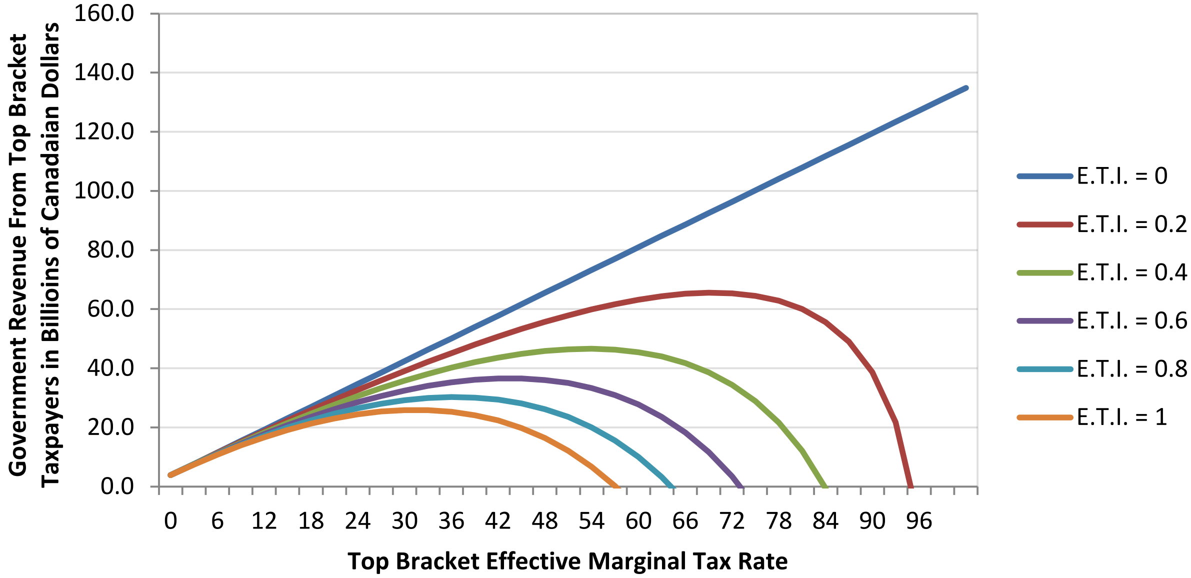
This section concludes with a collection of Laffer curves showing how total revenue collected from the top tax bracket changes with the top marginal tax rate under several different values of the E.T.I.

Since for E.T.I. = 0 we simply have the mechanical component of the increase in revenue from a change in top marginal tax rates, such Laffer curve are always linear.

For positive values of the E.T.I., the Laffer curve becomes the standard hill, although it is somewhat skewed to the right. It gets flatter and lower for higher values of the E.T.I., and the top of the curve, corresponding to the Laffer rate, moves to the left and is reached at a lower E.M.T.R.

Figure 9

Laffer Curves for Various Values of the E.T.I. - Canada



The last three figures, 10 to 12, show a direct comparison of Laffer curves across countries. By raising the Laffer rate, a country collects more tax revenue per taxpayer in the top tax bracket when measured in U.S. dollars and adjusted for purchasing power parity.

Figure 10

Comparison of Laffer Curves across Countries

E.T.I. = 0.2

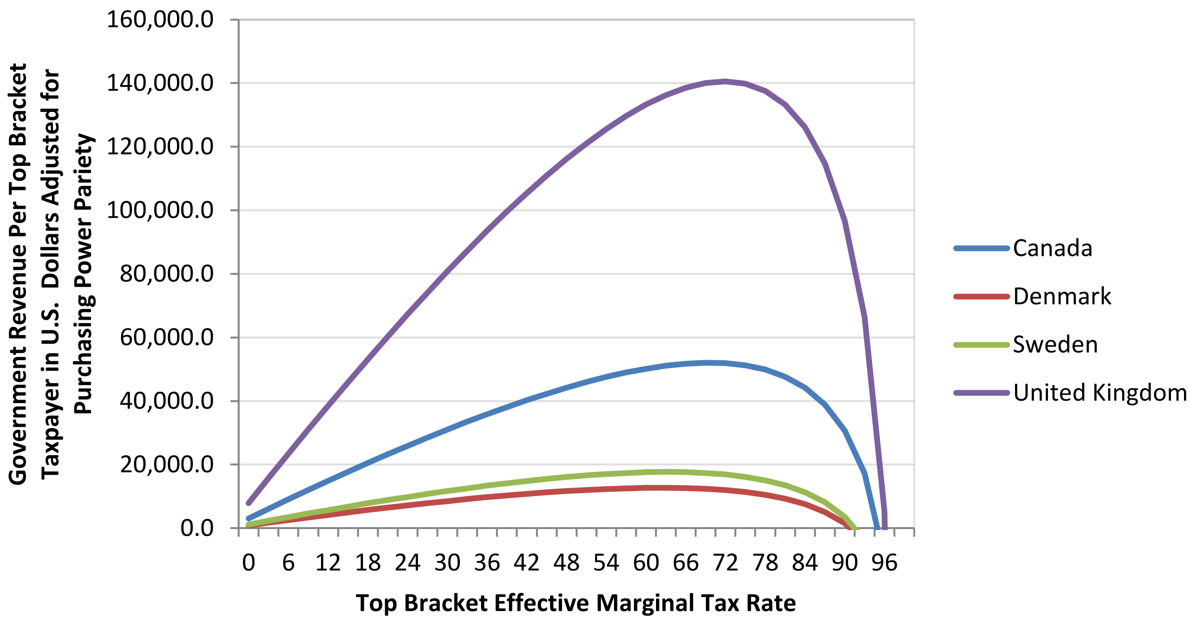


Figure 11

Comparison of Laffer Curves across Countries

E.T.I. = 0.4

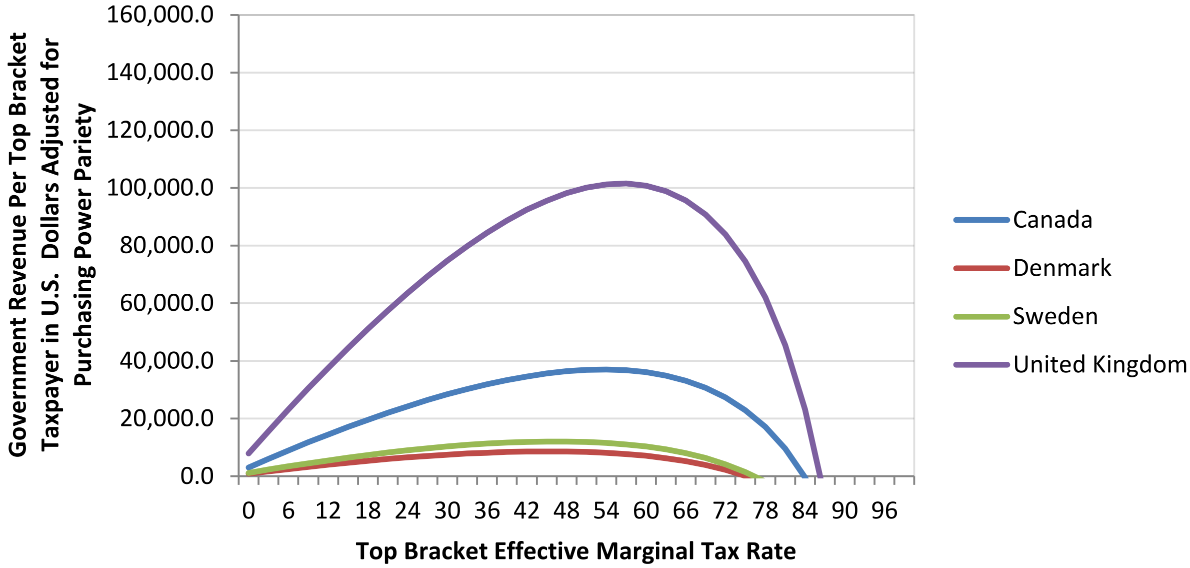
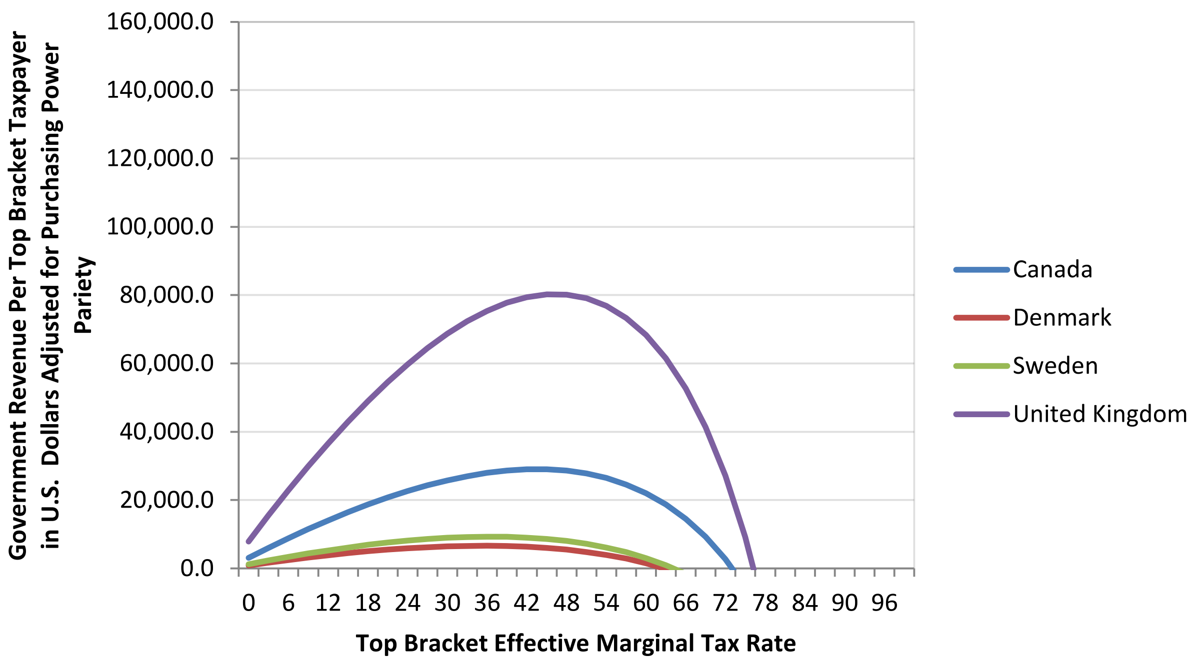


Figure 12

Comparison of Laffer Curves across Countries

E.T.I. = 0.6



4. WORKING WITH LESS DETAILED DATA

A) Introduction

The results presented thus far are exclusively for high income countries.

To examine issues related to raising tax revenue for low and middle income countries, a serious difficulty is limited availability of data.

Most income data on such countries are usually in the form of average incomes or income shares by grouped percentiles. Details of income distributions necessary to perform detailed studies on revenue collection and behavioural responses are not available.

This is particularly true of the highest income individuals. Inferences about their incomes cannot be meaningfully approximated from data on average incomes per population decile. Yet high income individuals present the largest behavioural responses to taxation, and are the source of a large portion of government revenue.

Getting around these limitations would allow us to obtain very interesting results, given the lower marginal tax rates generally in place in middle and low income nations.

For instance, Hourton (2012) provides income cut-off information per decile of the population in Chile using the Encuesta de Caracterización Socioeconómica Nacional (CASEN) for 2006. This is a mandatory survey issued by the government of Chile, and is the most reliable information on household income available to academics. However, since income data is based solely on amounts reported in the survey, inaccuracies relating to reporting errors by high income households present a serious challenge.

Putting such inaccuracies aside, once we have some information on income cut-offs in the income distribution, a feasible way to make inferences about those at the top of the income distribution is to look for ways to fit known income distribution functions to the limited data available.

One example was presented in Dagum (1977). He proposed a density function as follows:

$$f(y) = (1-\alpha)\lambda\beta\delta \bullet y^{-\delta-1}(1+\lambda \bullet y^{-\delta})^{-\beta-1} \text{ for income } y > 0 \quad (4)$$

To obtain the percentage of the population subject to the top tax rate, we use the Dagum CDF, given by:

$$F(y) = \alpha + (1-\alpha)/(1+\lambda \bullet y^{-\delta})^{\beta}, \lambda > 0 \quad (5)$$

Its inverse is given by:

$$y = \lambda^{1/\delta}\{[(1-\alpha)/(p-\alpha)]^{1/\beta}-1\}^{1/\delta} \quad (6)$$

where $p > \alpha$ is the probability of having an income less than or equal to y.

To obtain estimates of the average income for those in the top income tax bracket, we need to integrate the inverse of the CDF. The area of integration is from the percentage of the population subject to the top income tax rate to the top of the income distribution. Numerical estimation methods for the integral present one possible way forward.

An alternative approach which is explored in this thesis is developed in Saez, Slemrod & Giertz (2012). They recognize that the right tail of an income distribution is well approximated by a Pareto distribution. This has by definition a density function of the form $\boldsymbol{f(z) = C/z^{1+\alpha}}$. Denoting by $\boldsymbol{z^m}$ the average income of those in the top bracket and $\boldsymbol{\dot{z}}$ the income level where the top marginal tax rate kicks in, they point out that we have $\boldsymbol{z^m = \int zf(z)dz/\int f(z)dz}$, where the range of both integrals is from $\boldsymbol{\dot{z}}$ to infinity. This expression is in turn equal to $\boldsymbol{\dot{z}\bullet\alpha/(\alpha\text{-}1)}$, where in the results it is no longer necessary to know the value of $\boldsymbol{C}$. Hence, once the value of the Pareto α is known, assuming it is greater than 1 we can obtain an estimate for the average income for those in the top marginal tax bracket.

In order to obtain the required estimate of the Pareto α, we can look at a second functional form used to approximate income distributions, presented by Champernowne (1952). The density function is given by:

$$\boldsymbol{f(y) = (2N\alpha\sigma\sin\theta(y/y_0)^{\psi}/(y(1+\sigma)\theta\{1+2\cos\theta(y/y_0)^{\psi}+(y/y_0)^{2\psi}\})} \quad (7)$$

where $\psi = \alpha$ when $y > y_0$ and $\psi = \alpha\sigma$ otherwise.

Champernowne (1952) shows that the above parameter α converges asymptotically to the Pareto parameter α of the right tail of the income distribution in question, which is the α discussed just above. Hence, once we have an estimate of α for the Champernowne formula above, we can quickly calculate an estimate of the average income of those in the top tax bracket.

Note that the Dagum α has no direct connection to the Champernowne/Pareto α, and is designated by the same symbol by convention only.

In terms of techniques to estimate the parameters of the two income distribution functions given above, Campano (2006) provides an iterative program which selects a set of parameters for either the Dagum or Champernowne formulas so as to minimize the CHI square. The selection is taken from a collection of approximately 20,000 of the most likely combinations of such parameter values.

B) Revenue Forecasts Using Income Distribution Function Extrapolations

The goal of what follows is to investigate the potential of the Dagum, Champernowne and Pareto distributions in order to make inferences about underlying income distributions of nations that do not have a rich source of data readily available.

Returning to the case of Chile, we obtain the following:

Table 6

Income Distribution Results for Chile 2006

E.T.I. (Fernandez 2010)	2.0
Source of population data	Hourton (2012)
Dagum α	0.011312
Dagum β	1.3582
Dagum λ	0.67674
Dagum δ	1.6586
Champernowne α	1.5897
Champernowne θ	0.008422
Champernowne y_0	8911.3
Champernowne σ	1.2737
$\dot{z}$ (Income cut-off for top M.T.R.)	U.S. $124,100
Estimate of percentage of the population subject to top M.T.R. via Dagum C.D.F.	1.3773
N (number of people in the top tax bracket)	171,149
z (average income of those in the top tax bracket) via Pareto	U.S. $334,545
z via integration of Champernowne inverse C.D.F.	U.S. $311,743
Top marginal tax rate	40.00
% revenue lost to behavioural responses via Pareto distribution method	211.96
% revenue lost to behavioural responses via integration of the Champernowne inverse C.D.F.	221.52
Laffer rate via Pareto distribution method	23.92
Laffer rate via integration of the Champernowne inverse C.D.F.	23.13

The E.T.I. estimate comes from Fernandez (2010), who also uses the CASEN for 1996 to 2006 to obtain an estimate for Chile of 2.0. It is not possible to tell how much her estimates were influenced by reporting biases in the survey, as this estimate is very high.

Included in table 6 are estimates of the average income of those in the top tax bracket via integration of the Champernowne formula's inverse CDF function, as provided by Fred Campano.

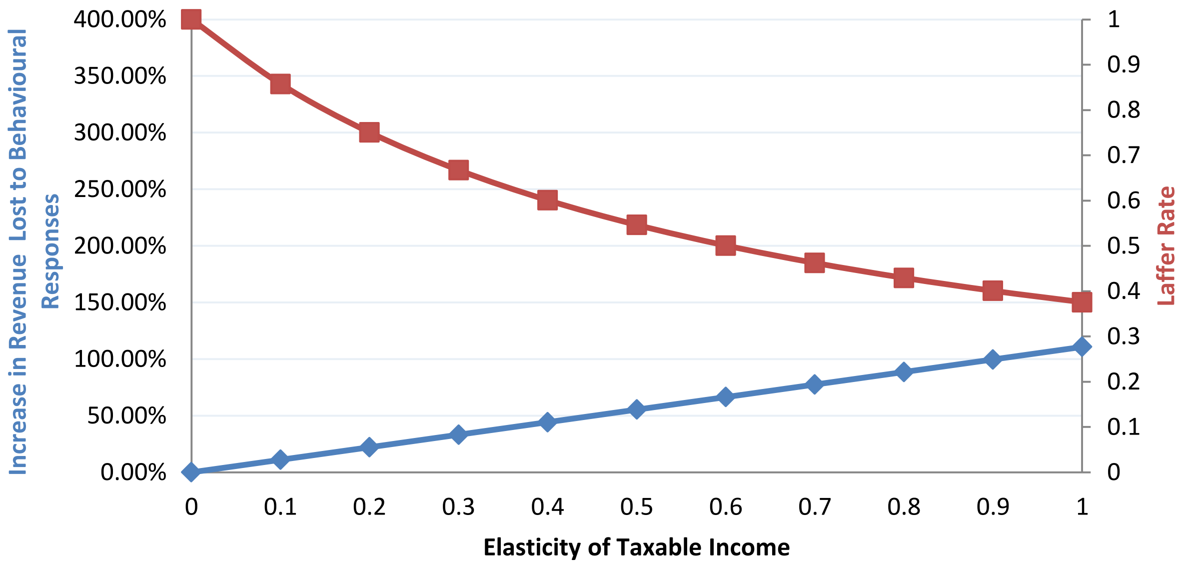
We can see that since the Pareto method may over-estimate top average incomes in this case, it provides a lower estimate of the potential percentage of revenue lost to behavioural changes.

The very high income cut-off for the top tax rate relative to average incomes heightens concern about the accuracy of the survey-based data. Since this is quite a small portion of the population, reporting errors by a subset of the group could have large impacts on both the E.T.I. estimate and the calculations of lost revenue.

Figure 13 summarizes potential lost revenue and Laffer rates for a range of more moderate E.T.I. values:

Figure 13

Potential Revenue Lost to Behavioral Responses & Estimates of Laffer Rates - Chile



Interestingly, potential revenue lost to behavioural responses for a given value of the E.T.I. is lower than many richer nations already studied. This is at least somewhat foreseeable due to the high average incomes of those in the top tax bracket and the discussion in section 3.

I now turn to Finland as an additional example of this method. Unlike Chile, Finland does not suffer from a lack of data generally. However, Statistics Finland only releases income distribution data in tabulated form by decile. Additionally, the W.T.I.D. (which was the source of income data in section 2) lacks detailed information on this country.

Mattika (2014) uses panel data involving 550,000 tax returns per year for the period 1995 to 2007 to obtain a plausible estimate of the Finnish E.T.I. of 0.475. The data cover the entire income distribution, and he presumes that the E.T.I. is applicable to all income groups, a strong assumption found frequently in the literature.

Using the Pareto method, we obtain the following results:

Table 7

Income Distribution Results for Finland 2007

E.T.I. (Mattika (2014))	0.475
Source of population data	Statistics Finland
Dagum α	0
Dagum β	0.87654
Dagum λ	445620
Dagum δ	4.0003
Champernowne α	4.0619
Champernowne θ	0.0988
Champernowne y_0	25467
Champernowne σ	0.88932
$\dot{z}$	€64,006.20
Percentage of population in the top tax bracket	2.268
N	99,610
z	€84,910.28
Top marginal tax rate	56.1
% revenue lost to behavioural responses	246.56
Laffer rate	34.14

Although the top marginal tax rate in Finland is comparable to other Nordic countries, the much higher E.T.I. estimate results in a very large amount of revenue predicted to be lost to behavioural responses.

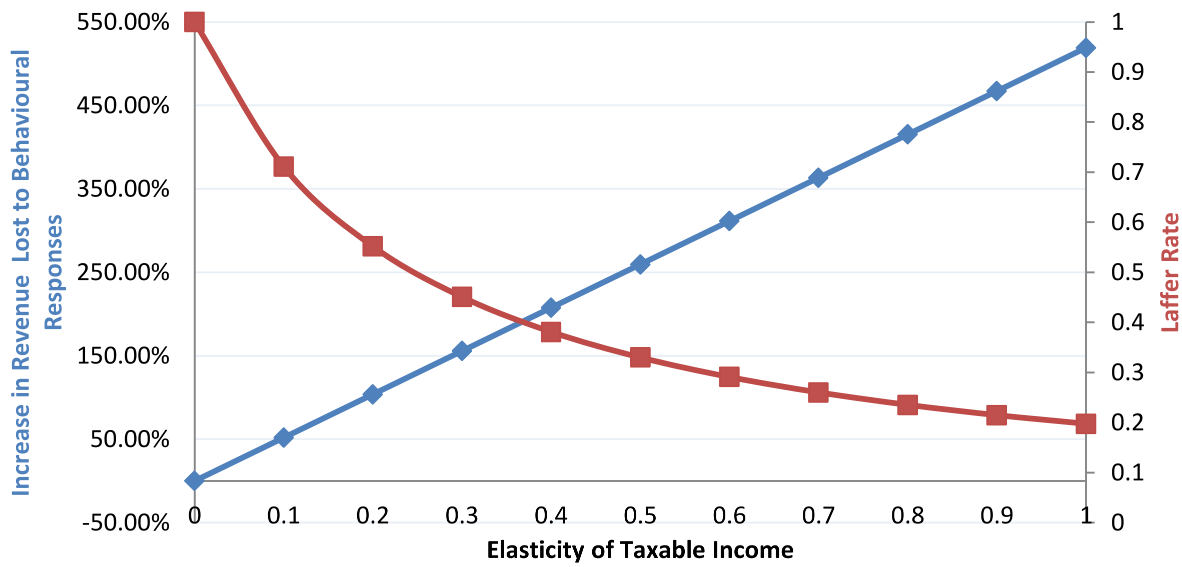
The higher value of the Pareto/Champernowne α relative to that of Chile fits with expectations that wealth inequality is not as extreme in Finland, as shown by a 'skinnier' right-hand tail of the income distribution.

The much smaller value of the Dagum α relative to Chile also gives less cause for concern in terms of the Dagum model's fit of the entire income distribution. The Dagum distribution function can be expected to give meaningful inferences over very nearly the entire income distribution.

Figure 14 presents Laffer rates and lost revenue for Finland. Noteworthy are the lower Laffer rates and higher percentages of lost revenue, consequences of a lower top bracket income cut-off and lower average income in that bracket.

Figure 14

Potential Revenue Lost to Behavioral Responses & Estimates of Laffer Rates – Finland



As a third example, the tax reform of New Zealand in 1990 makes for an interesting study. At that time, the government of New Zealand instituted a tax schedule which was very nearly flat, especially in comparison to other industrialized nations. The top tax rate was set to 33%. The level of income where the top rate kicked in was NZ$30,875, or just over US$18,000.

The study of New Zealand presented in section 3 focused on a period when the country had returned to a more progressive tax structure.

Although very detailed income distribution data is made publicly available by Statistics New Zealand, these go back only to 2001. Earlier years' data were never released.

However, Dixon (1996) uses the Household Economic Survey of New Zealand to tally income summaries for those who were working.

The Pareto distribution characteristics discussed above are assumed only to hold at the right tail of the income distribution, hence the technique is not applicable in this case. The Dagum formula predicts 31.8% of the population was subject to the top M.T.R., a group which is far from being constrained to that right-hand tail.

Instead, integrating the Champernowne formula's inverse CDF, we obtain the following the results:

Table 8

Income Distribution Results for New Zealand 1990

E.T.I. (Thomas (2007))	0.52
Source of population data	Dixon (1996)*
Dagum α	0
Dagum β	3.0967
Dagum λ	51.212
Dagum δ	1.7389
Champernowne α	2.0146
Champernowne Θ	143.24
Champernowne y_0	32558
Champernowne σ	0.49088
$\dot{z}$	NZ$30875
Percentage of population in the top tax bracket	31.800
N	829742
z	50299
Top marginal tax rate	33.0
% revenue lost to behavioural responses	66.32
Laffer rate	42.62

***Data only covers those who are employed**

Thomas (2007) used publicly unavailable panel data for the late 1980s and early 1990s to estimate an E.T.I. of 0.52., focused on labour income.

It is worth noting that using the estimate for the Pareto α to calculate the average income of those in the top income tax bracket gives a value of $61,305.72, an increase of over 20% compared to the integration method.

The estimated value of Pareto α obtained from the Champernowne α implies a more unequal wealth distribution than Finland, but less than Chile. This is in line with expectations. Additionally, large amounts of revenue are lost to behavioural responses with even a relatively modest estimate of 0.52 for the E.T.I.

Figure 15 presents Laffer rates and percentages of lost revenue for New Zealand during this period. The low value of the cut-off for the top income tax bracket allows us to predict the relatively low Laffer rates and percentages of lost revenue, again by way of the discussion in section 3.

The remainder of the figures in this section show the relationship between Chile and richer nations in terms of potential lost revenue and Laffer rates.

Specifically, figures 16 and 17 compare potential lost revenue and Laffer rates across countries, 18 looks at revenue per top-bracket taxpayer for the case of Chile, and 19 and 20 compare such revenue per top-bracket tax payer across countries. Tables 9 and 10 examine which are the driving forces behind variations across countries.

Figure 15

Potential Revenue Lost to Behavioral Responses & Estimates of Laffer Rates - New Zealand 1990

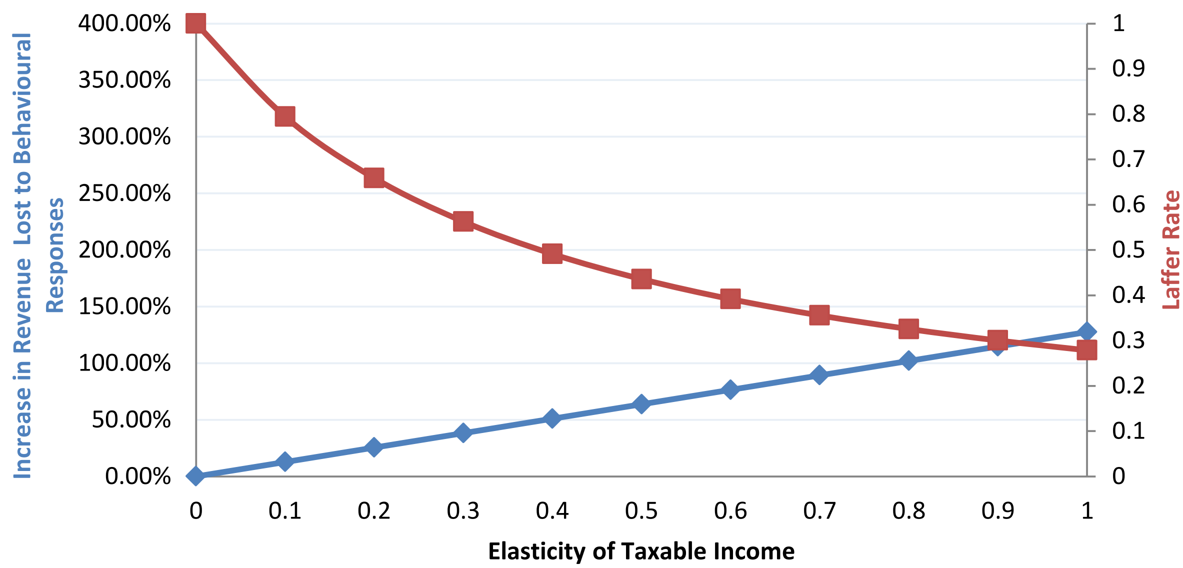


Figure 16

Comparison of Potential Revenue Lost to Behavioral Reactions across Countries

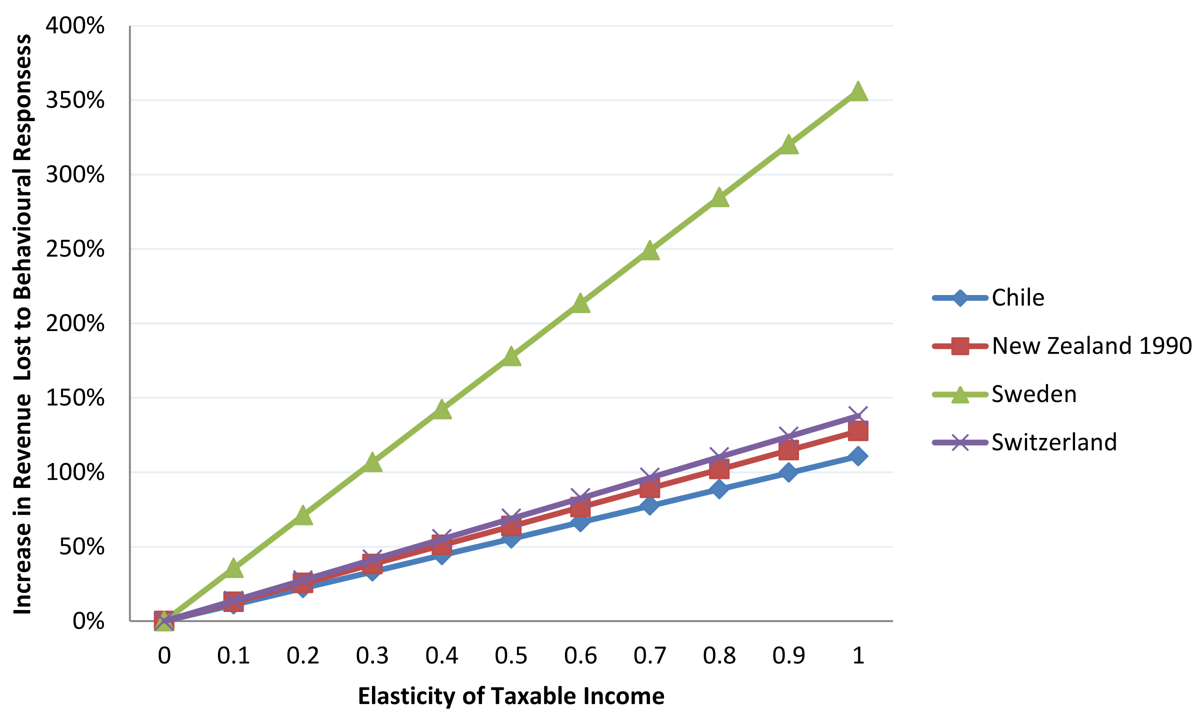


Table 9

Partial Derivatives of the Equation for Percentage of Revenue Lost to Behavioural Responses
Taken With Respect to Its Influencing Factors
Evaluated at Most Current Values of Each Variable

Partial Derivative W.R.T.	Chile	New Zealand 1990	Sweden	Switzerland
z	-2.80×10^{-05}	-2.20×10^{-05}	-5.70×10^{-05}	-3.50×10^{-06}
ż	3.73×10^{-05}	3.66×10^{-05}	9.01×10^{-05}	7.26×10^{-06}
EMTR	11.53	5.49	14.50	5.68

Sweden is again highest in all values. The fact that Chile and New Zealand have such large differences in the category for marginal tax rates for top earners again suggests that average incomes in the top tax bracket may be very influential in how much potential revenue is lost to behavioural responses.

Table 10

Ratios of Partial Derivatives of the Equation for Percentage of Revenue Lost to Behavioural Responses
Taken With Respect to Its Influencing Factors
Evaluated at Most Current Values of Each Variable

	Chile	New Zealand 1990	Sweden	Switzerland
ż/z	-1.31	-1.68	-1.58	-2.05
EMTR/z	-405029	-252474	-254452	-1604616
EMTR/ ż	309067	149945	160824	782002

Figure 17

Comparison of Laffer Rates Across Countries - Restriction to Higher Values of the E.T.I.

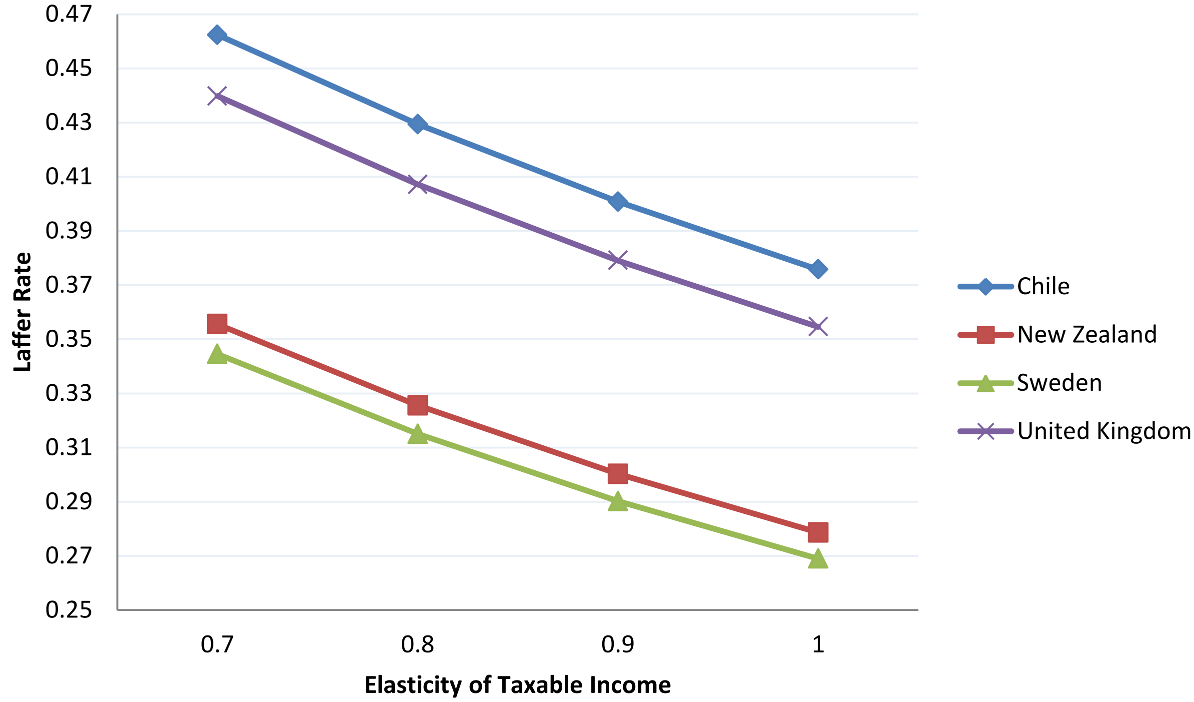


Figure 18

Laffer Curves for Various Values of the E.T.I. - Chile

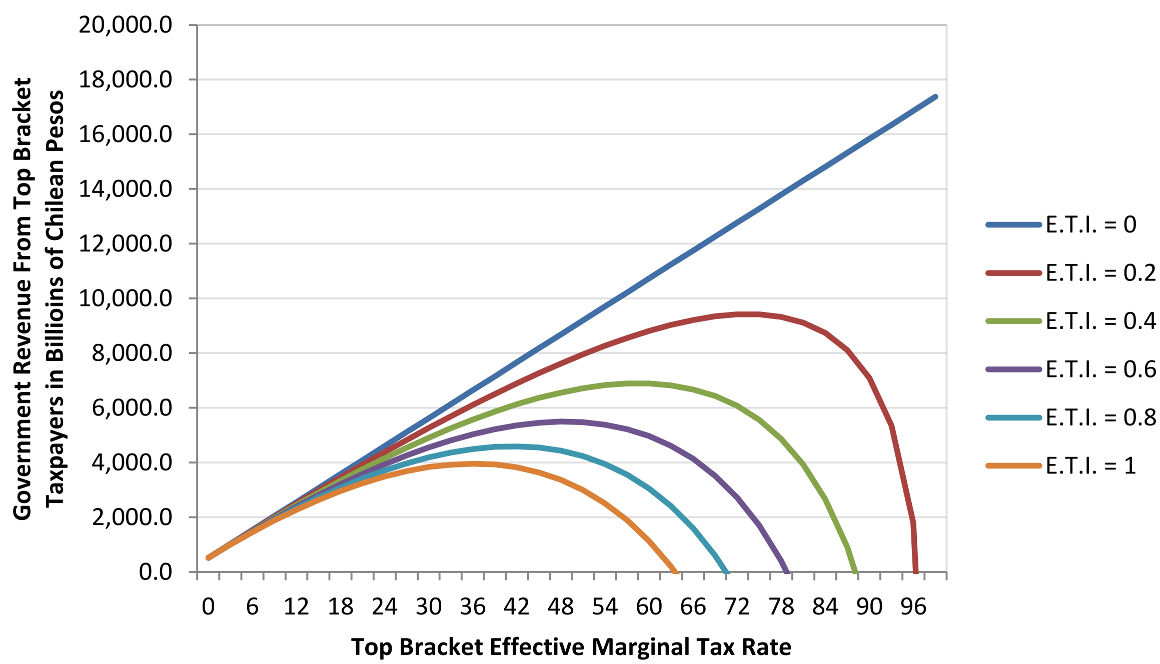


Figure 19

Comparison of Laffer Curves across Countries

E.T.I. = 0.2

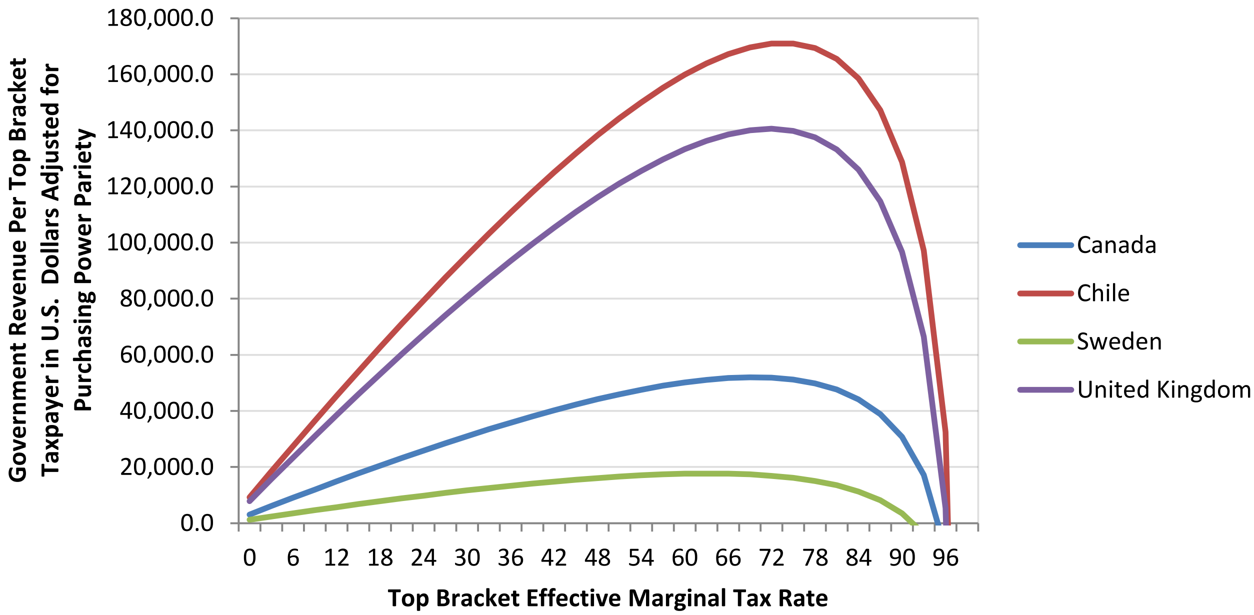
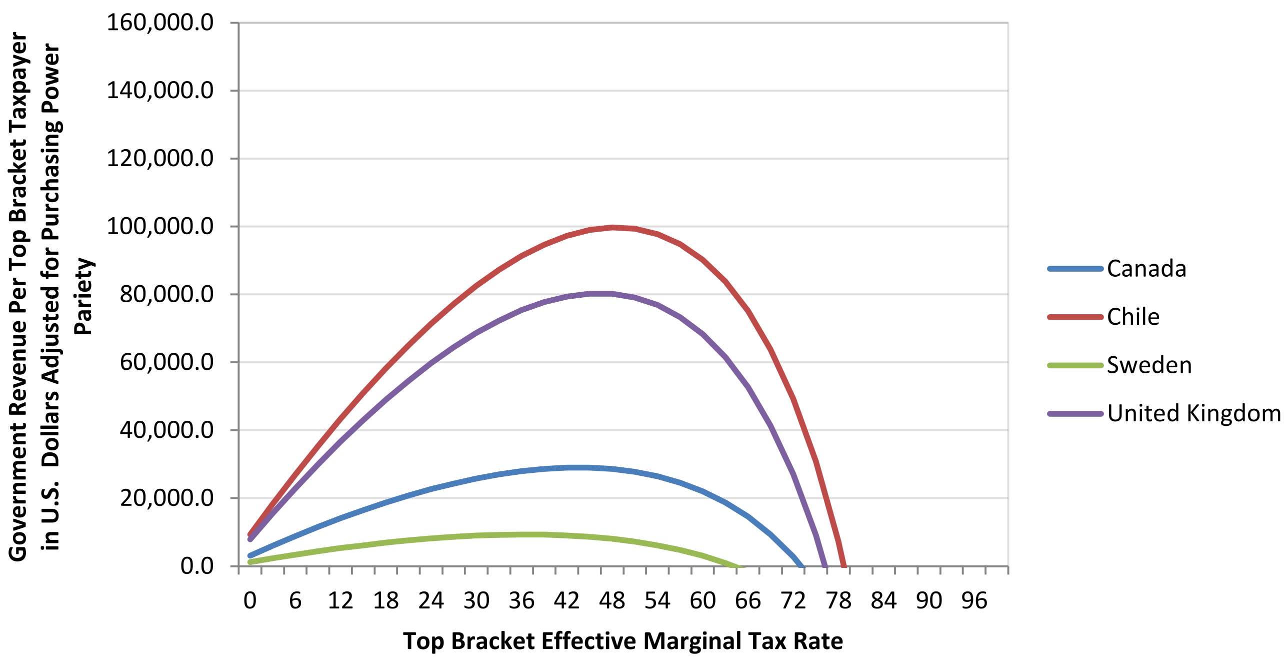


Figure 20

Comparison of Laffer Curves across Countries

E.T.I. = 0.6



C) Testing Goodness of Fit

As a final point on this topic, I turn to the issue of how well the Dagum and Champernowne functions can be expected to fit the true income distributions.

Although Campano (2006) provides some analysis on this issue, comparing the fit of these distributions at the national level across countries is not one of his focal points.

Because of the lack of a central source for income distribution data, it is quite difficult to find information on both decile income cut-offs such as those used above to estimate the Dagum and Champernowne parameters, and detailed data on incomes in the right-hand tail of income distributions. Together, these two sets of data could be used to check how well the Dagum and Champernowne functions estimated with income data by population decile fit the true distribution of the right hand tail.

The income data used in previous sections to obtain direct estimates of revenue lost to behavioural responses were usually obtained through the World Top Incomes Database, which does not provide detailed information about the population outside of the top 10% of earners.

I have, however, found both types of data for certain years for Canada, Germany and New Zealand.

The data provides information on average incomes and income cut-offs for various percentiles of the population.

I begin by restricting the data to that of population deciles, and estimate corresponding parameter values for the Dagum and Champernowne distributions.

Table 11
Income Distribution Function Parameter Estimates
Testing for Goodness of Fit

Country	Canada	Germany	New Zealand
Year	2009	1998	2011
Currency used	Cdn $	Euros	NZ $
Source of data	Hunter, Sanchez & Douglas (2012)	Bach, Corneo & Steiner (2006)	Statistics New Zealand
Dagum α	$2.3965X10^{-9}$	0.053145	$1.0694X10^{-7}$
Dagum β	0.57937	0.21045	0.28513
Dagum λ	156.55	5182.5	161.4
Dagum δ	2.4957	5.2898	2.9703
Champernowne θ	$7.9591X10^{-12}$	161.12	$9.6338X10^{-11}$
Champernowne α	2.6513	0.90685	3.4603
Champernowne y_0	6956.7	37236	49104
Champernowne σ	0.61288	0.5453	0.30254

After obtaining the parameter estimates, I use income cut-offs available from the data and the Dagum income distribution to predict the percentage of the population with incomes above that cut-off (denoted 'Predicted %' in the preceding tables). I then compare the results to the true percentage of the population with income above the cut-off according to the data (denoted 'Observed %' in the preceding tables). 'Difference' refers to the predicted percentage minus the observed percentage.

Next, I use the technique discussed previously and the estimate of Pareto α from the Champernowne distribution to estimate average incomes for those in the observed percentiles just analysed (denoted 'Predicted income average'). The corresponding true average incomes are available from the data. I then divide the predicted average income by the observed average income of the data, putting it into percentage form. The column 'Percent difference' is simply the result subtracted from 100.

Negative estimates are explained shortly after the tables presenting the results.

The results are a test of how well the predictions from the above techniques could be expected to fit the true underlying values of $\boldsymbol{z}$ and $\boldsymbol{N}$.

Table 12

Goodness of Fit Results for Canada 2009

(Figures in Canadian Dollars)

Income cut-off	Observed %	Predicted %	Difference	Observed income average	Predicted income average	Percent difference
140,000	10.00	10.71	0.71	248,110	224,781.69	9.40
179,400	5.00	6.17	1.17	335,475	288,041.68	14.13
301,975	1.00	1.79	0.79	647,454	484,846.07	25.11

Table 13

Goodness of Fit Results for Germany 1998

(Figures in Euros)

Income cut-off	Observed %	Predicted %	Difference	Observed income average	Predicted income average	Percent difference
64,900	10.00	4.52	-5.48	116,050	-631,826	644.44
83,600	5.00	1.31	-3.69	159,300	-813,877	610.91
162,700	1.00	0.04	-0.96	367,700	-1,583,945	530.77
551,900	0.10	0.00	-0.10	1,435,100	-5,372,952	474.40

Table 14

Goodness of Fit Results for New Zealand 2011

(Figures in New Zealand Dollars)

Income cut-off	Observed %	Predicted %	Difference	Observed income average	Predicted income average	Percent difference
72,000	9.97	10.20	0.23	117,493.70	101,264.72	13.81
75,000	8.96	9.26	0.30	122,466.90	105,484.09	13.87
78,000	8.02	8.42	0.40	127,818.00	109,703.45	14.17
82,000	6.98	7.44	0.46	134,945.80	115,329.27	14.54
87,000	5.94	6.41	0.47	143,789.10	122,361.54	14.90
93,000	4.95	5.39	0.44	154,544.00	130,800.27	15.36
101,000	3.98	4.33	0.35	168,722.90	142,051.90	15.81
113,000	2.99	3.18	0.19	189,278.90	158,929.36	16.03
131,000	2.01	2.10	0.09	222,135.00	184,245.54	17.06
170,000	1.00	1.00	0.00	297,006.10	239,097.26	19.50

Note: Data from Statistics New Zealand is grouped by amount of income in thousand N.Z. dollar increments. This approach differs from the standard grouping by population top-income percentiles. Hence the percentages of the population used differ somewhat from the usual round percentiles found elsewhere. More percentiles are examined because of the greater detail of the data compared to the sources available for Canada and Germany.

Generally, the Dagum model predicts the percentile corresponding to a given income reasonably well for Canada and New Zealand, usually within 1 point, and often much less.

For these two countries, the fit of the estimates of average income within the percentile is poorer. Close to the top 10th percentile the fit tends to be somewhat better, but we notice a general deterioration as we move farther to the right in the distribution. The predictions are consistently lower than the observed value, which means that the amount of revenue lost predicted with these methods may be lower than what would truly be observed.

The models fail to give useful predictions for Germany, in large part because the estimate for the Pareto α is less than 1. Since we use the formula $\boldsymbol{z^m = \dot{z} \bullet \alpha/(\alpha\text{-}1)}$ to calculate average income, we obtain negative values, which have no economic meaning.

In summary, the preceding examples show that where the coefficients are reasonable (most obviously that Pareto $\alpha > 1$), the models can provide useful information. They give at least rough approximations to the average incomes of those in the right tail of the income distribution. However, to obtain closer estimates of average incomes, it is better to estimate the model based on data from farther to the right of the income distribution than the 10th percentile. Unfortunately, many data sources do not enable us to do so.

At the very least, using the techniques outlined in this section allows us a general examination of revenue potential in a wider range of countries.

5. THE IMPACT OF EXCLUDING CAPITAL GAINS FROM INCOME MEASURES

A) Introduction

A possible confounding factor in estimation techniques of the E.T.I. discussed in section 2 is the common practice of omitting capital gains from definitions of income (whether capital gains are included in the definition of income used to calculated effective marginal tax rates is not addressed in articles dealing with this topic). This income definition is then used in time series or panel data regressions to estimate elasticities.

If the responsiveness of capital gains realizations to changes in tax rates is different than that of other sources of income, the resulting measure of the E.T.I. will be biased. When these E.T.I. estimates are used to estimate changes to revenue via changes in marginal tax rates, the results will also be inaccurate. The problem becomes more serious the greater the percentage of national income generated via capital gains.

The problem is all the more acute when the tax reform used for identification contains adjustments specific to the taxation of capital gains. As an example, the major tax reforms in the U.S. during the 1980s repealed a 60% exemption on capital gains realizations, while significantly lowering marginal tax rates for income in general. This means that the taxation of capital gains increased substantially relative to that of other income such as dividends. As a result, there was a large incentive to change the focus of investment portfolios from growth stocks to dividends, something which is not difficult for investors to do when institutional investors may not have the same tax incentives.

Yet the definition of taxable income used in the E.T.I. studies of that period (Feldstein (1995), Auten and Carroll (1999), Gruber and Saez (2002), Saez (2004), etc.) does not include capital gains. It does however incorporate dividends. In the case of the 1986 tax reforms, post-reform increases in declared dividends will appear as increases in income via behavioural responses to lower tax rates. If without the tax reform such income would have instead come from capital gains, income shifting is misrepresented as income growth, and estimates of the E.T.I. will be biased upward.

Auten and Carroll (1999) examine the statistical significance of dummy variables for having realized capital gains, as well as being subject to alternative minimum tax. They conclude that since they chose not to include those subject to alternative minimum tax in the data sample used for their preferred E.T.I. estimates, any bias caused by the exclusion of capital gains from their definition of income will be minimal. Their logic is that those with large capital gains realizations are subject to alternative minimum tax, and hence capital gains do not figure prominently in income declared by those in their samples.

Due to its simplicity, this argument is far from conclusive. Not distinguishing between those with capital gains of a few dollars as opposed to those with significant portions of their revenue coming from capital gains could be problematic. The latter are likely to have higher incomes and much larger incentives to adjust their portfolios. Still, capital gains may not be a high enough percentage of their income for them to be subject to alternative minimum tax.

Likewise, the self-employed often have the option to pay themselves in dividends or by selling stock in their companies. Executives paid in stock options also have to report capital gains. Yet such individuals will not necessarily be subject to alternative minimum tax calculations.

Even assuming that the present value of all future capital gains realizations will be constant, significant inter-temporal shifting of capital gains to minimize tax liability could have significant impacts on the present value of all future government revenue.

Further investigation of this issue would be very useful, yet because the taxation of capital gains usually changes concurrently with that of other income, identification is challenging.

However, starting in the late 1980s in Canada, there were a series of changes to the inclusion rate of capital gains (the percentage of capital gains income which must be declared as income) which were not accompanied by changes to the federal marginal tax rate. This scenario offers a valuable opportunity to isolate the response of individuals to changes in the taxation of capital gains.

My approach is to perform a simple time series regression of total capital gains realizations in Canada on the inclusion rate of capital gains, controlling for the performance of the stock market.

This technique follows such work as that of Saez (2004) and Milligan & Smart (2013).

The regression model is as follows:

$$\mathbf{cgr}_t = \alpha + \beta\mathbf{inc}_t + \delta_1\Delta\mathbf{stock}_t + \delta_2 t + \delta_3 t^2 + \varepsilon_t \qquad (8)$$

Where:
cgr_t denotes capital gains realized in year t.
α is the regression constant.
inc_t is the inclusion rate in year t.
$\Delta stock_t$ is the percentage change in the stock market in year t.
ε_t is the error term of zero conditional expected value.

The models are estimated with and without linear and quadratic time trend controls, which can account for changes to investment strategies independent of taxes.

The period covered is from 1972, when capital gains became taxable, to 2013, the latest year for which data are available.

The issue of lagged responses to changes in the inclusion rate is discussed shortly.

My preferred model has cgr_t and inc_t in logarithmic form since β becomes the elasticity of capital gains realizations in terms of the inclusion rate.

Using Durbin-Watson and Dickey-Fuller tests suggests significant autocorrelation in the error terms, so Newey-West standard errors with 3 lags are presented in the results. The choice of 3 lags is based on the standard formula $g = 4(n/100)^{2/9}$, where n is the sample size and the number of lags is chosen to be the highest integer less than g (Newey & West (1987)). Hence 3 lags are optimal when the sample size is around 50.

In what follows, I begin by presenting the available data, before turning to my results and several alternative specification models. I conclude this section with a look at capital gains taxation across various countries in the O.E.C.D.

B) Data

Data on capital gains realizations and marginal tax rates come from the annual Taxation Statistics publication of Revenue Canada. Data on other types of investment income come from Statistics Canada CANSIM tables 376-0013, 376-0012, and 376-0136. Information on the performance of the Canadian stock market comes from the O.E.C.D. Financial Statistics database.

Capital gains were not taxed before 1972, at which point the inclusion rate was set at ½. It remained at this level until 1988, when it was raised to ⅔, before rising again to ¾ in 1990. It stayed at this level until 2000, when it was progressively lowered throughout the year, with an average rate for 2000 of ⅔. In 2001 it was lowered to ½, where it has remained since.

Unfortunately, records of capital gains realized prior to 1972 are not publicly available, meaning that my model is unable to take advantage of a very interesting potential source of identification - the introduction of capital gains taxation in 1972.

The federal top marginal tax rate has been reasonably stable since 1972, changing just three times: first, from 47% to 43% in 1976, then to 34% in 1982 and finally to 29% in 1988, where it stayed until 2016. The changes of 1988 coincided with changes to the inclusion rate of capital gains, but the changes to the M.T.R. were minor in comparison to that of the inclusion rate. Certainly, the U.S. M.T.R. changes of the 1980s were striking in comparison (from 50% to 28% on top incomes for the changes of 1986).

C) Results

First, the primary factor affecting capital gains realizations is the performance of the stock market. The two factors follow each other quite closely, as is illustrated by figures 21 and 22:

Figure 21

Capital Gains Realizations and the Value of the Stock Market (Nominal Values)

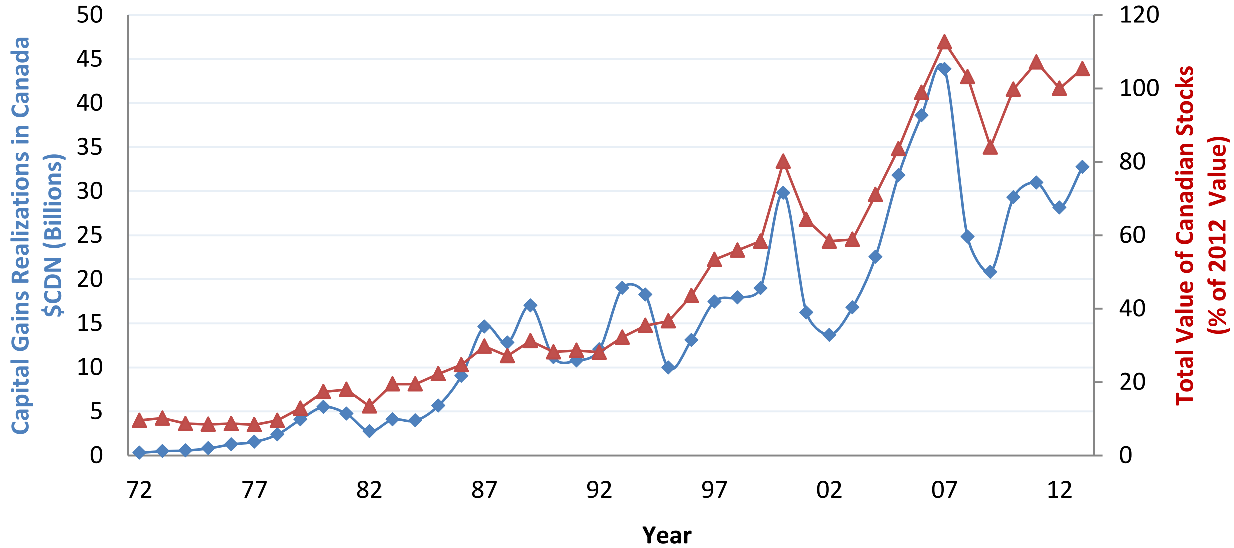
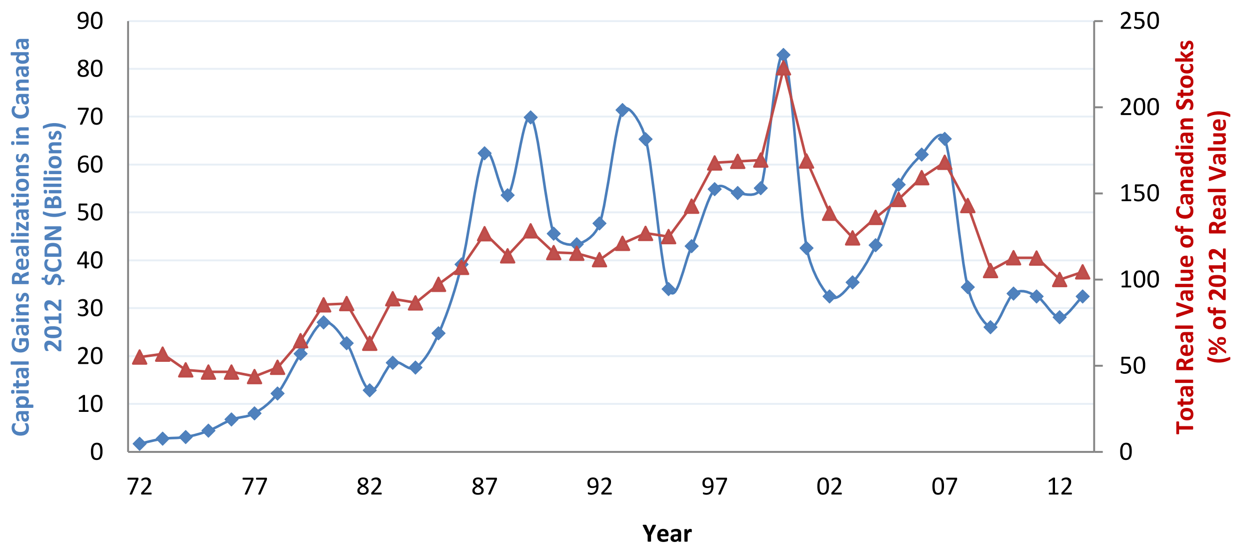


Figure 22

Capital Gains Realizations and the Value of the Stock Market (Real Values)

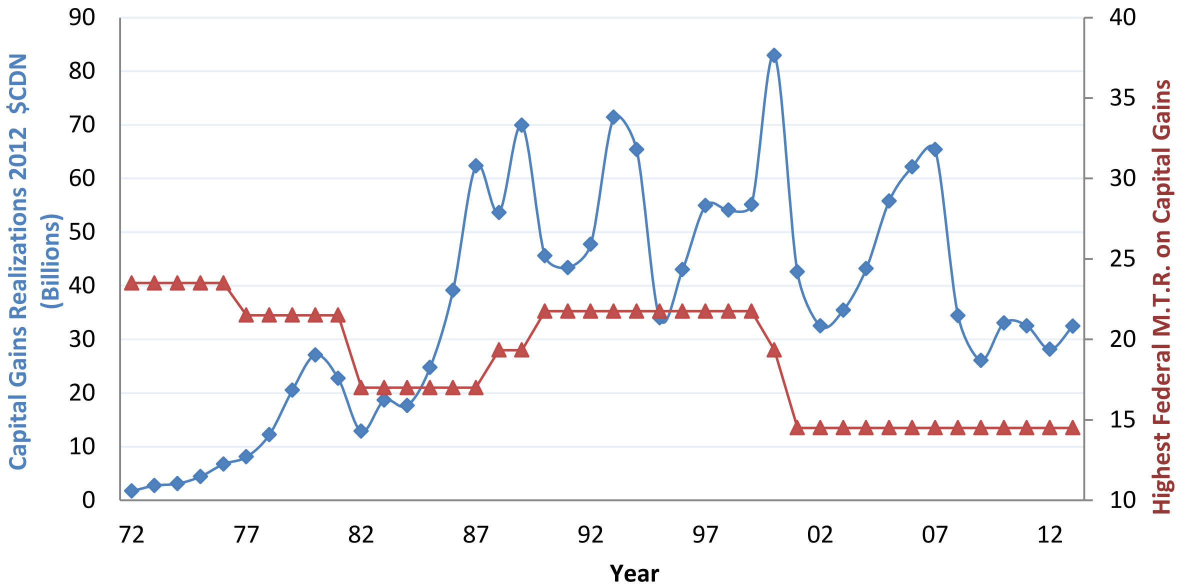


Therefore in all regressions that follow, a control is included for the annual performance of the stock market based on data supplied by the O.E.C.D.

Figure 23 shows the relationship between capital gains realisations and federal effective M.T.R.s on capital gains, respectively. It is noticeable that capital gains realizations on the federal and Ontario levels follow similar paths, but on a different scale. No immediate visual relationship is noticeable between capital gains realizations and M.T.R.s in either case:

Figure 23

National Capital Gains Realisations and Top Federal M.T.R. on Capital Gains (in real terms)



Note that in figure 23 top marginal tax rates on capital gains refer to tax obligations including both the inclusion rate and marginal tax rate on taxable income.

Figures 24 and 25 show changes in nominal amounts of investment income, broken down by source (interest income on bonds, dividend income, and capital gains income). Even in nominal terms, we see that interest income has been relatively flat, while both capital gains and dividends income have increased substantially. This can be seen as justification for including time trends in the regression.

Figure 24

Stacked Breakdown of Components of Investment Income (Nominal Terms)

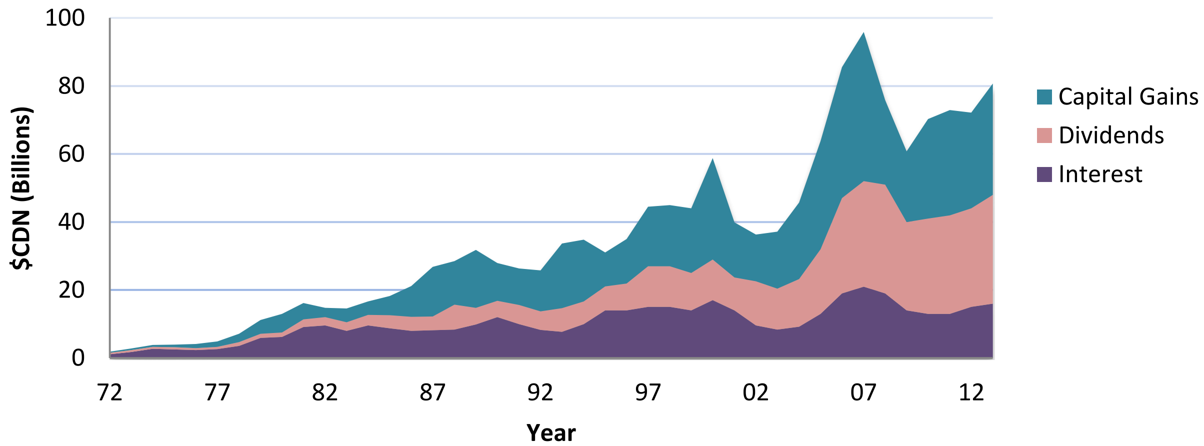


Figure 25

Stacked Breakdown of Components of Investment Income (Real Terms)

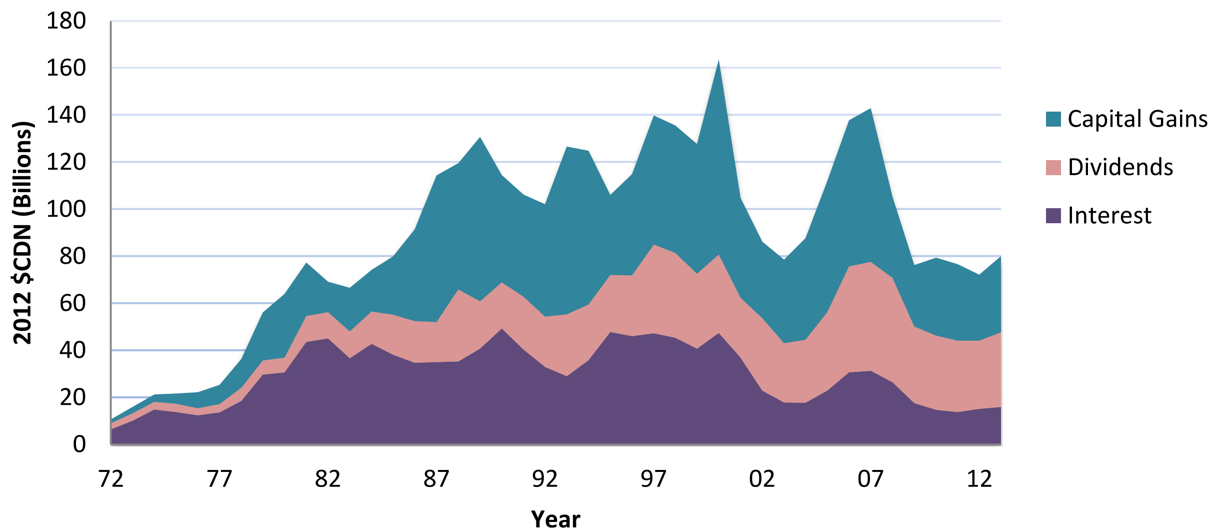
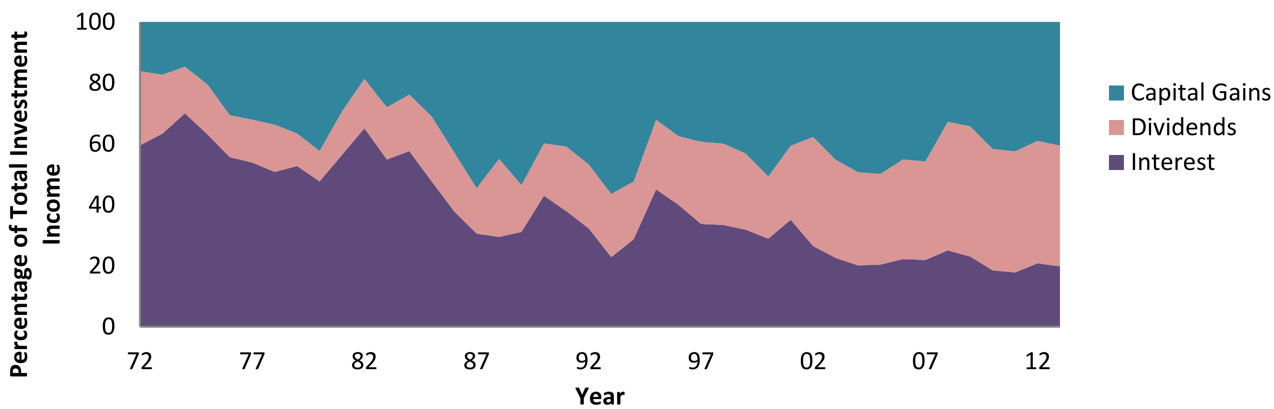


Figure 26 divides investment income per year into its components on a percentage basis, where the shift to dividend income is striking. Although this has come primarily at the expense of interest income, the portion of capital gains decreases slightly as that of dividends grows, especially since the turn of the century. This is further justification for including time trends.

Figure 26

Components of Investment Income in Percentages of Total



Turning to the results of the regressions, in table 15 we see that unless we include a polynomial time trend, capital gains realizations are predicted to rise with the taxation of such income, which is counter-intuitive.

Table 15

Regression in Log Form of Capital Gains Realizations on the Inclusion Rate – Canada
Standard Errors: Newey-West

	Regression in Logs	Regression in Logs with Linear Time Controls	Regression in Logs with Polynomial Time Controls
Coefficient	1.95	1.13	-0.59
Standard Error	1.38	0.47	0.50
t-statistic	1.41	2.41	-1.19

We see similar results if we lag the inclusion rate by one year, which is to assume that changes to investment strategies take a year to become apparent.

Table 16

Regression in Log Form of Capital Gains Realizations on a One Period Lag of the Inclusion Rate – Canada

Standard Errors: Newey-West

	Regression in Logs	Regression in Logs with Linear Time Controls	Regression in Logs with Polynomial Time Controls
Coefficient	1.79	0.83	-0.70
Standard Error	1.29	0.39	0.48
t-statistic	1.38	2.12	-1.45

Turning now to dividends, I examine the relationship of total dividends received by Canadians in the year with the inclusion rate on capital gains. Without time trend controls, the estimates are of the expected sign, but not statistically different from zero. The regression with a linear time control results in a positive and statistically significant estimate of the relationship between dividends and the inclusion rate in the year of the change. The elasticity becomes negative when a quadratic time control is included, although with a very low t-statistic.

Table 17

Regression in Log Form of Received Dividends on the Inclusion Rate – Canada

Standard Errors: Newey-West

	Regression in Logs	Regression in Logs with Linear Time Controls	Regression in Logs with Polynomial Time Controls
Coefficient	1.05	0.62	-0.21
Standard Error	1.46	0.24	0.36
t-statistic	0.72	2.56	-0.57

As a first robustness check, I now turn to estimates of the coefficients on the elasticities of dividends and capital gains as before, but which now include up to 5 lags in the same regression model. I also incorporate the method of Almon polynomial distributed lag models (Almon (1965)). Quadratic and cubic polynomials are used, with 3 and 5 lags. Controls for time trends and the performance of the stock market are also included.

Perhaps due to the small number of degrees of freedom, most estimates of the coefficients are not statistically significant.

We generally notice little difference in the estimates of direct models versus those using a polynomial technique. This is not entirely surprising, since even a degree 2 polynomial requires 3 control variables, creating an advantage of just 2 degrees of freedom when using 5 lag models, and an advantage of just 1 degree of freedom when a 3 lag model is used.

Table 18

Regression in Log Form of Capital Gains Realizations on Various Lags of the Inclusion Rate – Canada
Polynomial Distributed Lag Model with Quadratic Time Controls
Standard Errors: Newey-West

Years of Lag of the Inclusion Rate	Regular Regression with 3 Lags	Regular Regression with 5 Lags	Degree 2 Polynomial with 3 Lags	Degree 2 Polynomial with 5 Lags	Degree 3 Polynomial with 3 Lags	Degree 3 Polynomial with 5 Lags
0	0.16 (0.66)	0.13 (0.75)	0.19 (0.64)	0.18 (0.42)	0.16 (0.66)	0.19 (0.68)
1	0.39 (0.57)	0.20 (0.69)	0.32 (0.43)	0.05 (0.10)	0.39 (0.57)	0.04 (0.47)
2	-0.22 (0.65)	-0.25 (0.57)	-0.15 (0.53)	-0.10 (0.26)	-0.22 (0.65)	-0.10 (0.46)
3	-1.21 (0.40)	-0.20 (0.44)	-1.23 (0.36)	-0.24 (0.26)	-1.21 (0.40)	-0.23 (0.17)
4		-0.32 (0.50)		-0.39 (0.11)		-0.37 (0.33)
5		-0.58 (0.37)		-0.54 (0.39)		-0.55 (0.41)

Returning to dividends, we notice an increase in dividends reported in the year of the change in inclusion rates, and this is often significant at even the 1% level. However, we often see a large swing in the opposite direction in the year following the change in the inclusion rate. It is not immediately clear what could explain this.

A possible explanation is that in the 1990s (where most of the variation in inclusion rates used for identification takes place) there was a general trend by companies to reduce the payouts of dividends (Graham (2005)). A summary of dividend payouts is presented in the appendix.

Table 19

Regression in Log Form of Received Dividends on Various Lags of the Inclusion Rate – Canada
Polynomial Distributed Lag Model with Linear Time Controls
Standard Errors: Newey-West

Years of Lag of the Inclusion Rate	Regular Regression with 3 Lags	Regular Regression with 3 Lags	Degree 2 Polynomial with 3 Lags	Degree 2 Polynomial with 5 Lags	Degree 3 Polynomial with 3 Lags	Degree 3 Polynomial with 5 Lags
0	1.75 (0.63)	1.57 (0.54)	1.47 (0.64)	0.82 (0.31)	1.75 (0.63)	1.17 (0.56)
1	-1.17 (0.54)	-1.18 (0.48)	-0.40 (0.51)	0.13 (0.11)	-1.17 (0.54)	-0.29 (0.46)
2	-0.07 (0.81)	-0.07 (0.68)	-0.83 (0.55)	-0.27 (0.23)	-0.07 (0.81)	-0.51 (0.41)
3	-0.11 (0.60)	0.51 (0.37)	0.18 (0.50)	-0.40 (0.21)	-0.11 (0.60)	-0.14 (0.17)
4		-0.78 (0.33)		-0.24 (0.08)		0.18 (0.39)
5		0.20 (0.32)		0.19 (0.38)		-0.19 (0.36)

As a final robustness check, it is plausible that changes to the amounts of capital gains declared may take some time to appear, as the average investor may hold onto an investment for a certain period.

It is estimated that the average holding period for an investment in the period from 1972 to 2000 was anywhere from 2 to 5 years (Dela Croce, Stewart & Yermo (2011), Kleintop (2012), Haldane (2010)). Therefore, I repeat the regression of capital gains realisations on the inclusion rate as in equation (8), but with a specific lag on the inclusion rate.

Table 20 presents the results.

Table 20

Regression in Log Form of Capital Gains Realizations on Lags of Various Periods of the Inclusion Rate – Canada

Standard Errors: Newey-West

Years of Lag of Inclusion Rate	Regression in Logs	Regression in Logs with Linear Time Controls	Regression in Logs with Polynomial Time Controls
2	1.64 (1.26)	0.56 (0.38)	-0.92 (0.45)
3	1.51 (1.19)	0.30 (0.36)	-1.06 (0.35)
4	1.44 (1.09)	0.09 (0.35)	-1.09 (0.31)
5	1.50 (0.98)	-0.03 (0.34)	-1.05 (0.32)

Using a polynomial time trend control, we see that the estimates of the elasticity of capital gains realizations to the inclusion rate are now quite large, of the expected sign, and statistically significant at even the 1% level. Importantly, they are an order of magnitude higher than standard estimates of the E.T.I.

Without a polynomial time trend control, estimates are not statistically different from zero at any conventional significance levels, however.

The idea that tax considerations do not impact investment decisions is quite counter-intuitive, hence we may expect the quadratic time control model best fits the underlying relationship.

Importantly, if this is true, we notice a very large elasticity of capital gains realizations to the corresponding tax rate. These results illustrate that excluding capital gains from measurements of income could significantly bias estimates of the E.T.I. downwards.

Taking a moderate estimate of the E.T.I. for Canada such as 0.4 (Milligan & Smart (2014), and returning to figure 2, we see that even a 0.1 increase in the E.T.I. estimate is associated with more than a 15 point increase in the percent of revenue gains lost to behavioural responses. Since upper-income individuals often have substantial percentages of their income derived from capital gains, the revenue-forecasting implications of this bias are quite alarming.

To conclude this section, table 18 shows how many countries currently have significant tax incentives to focus portfolios toward capital gains. It shows the top effective marginal tax rate on capital gains in a particular country minus that on dividends. That is, the point difference between the two.

E.T.I. estimates from those countries with large differences between effective marginal tax rates on capital gains and dividends could be especially susceptible to bias caused by the exclusion of capital gains from estimates of taxable income.

Table 21

Differences between Top Marginal Tax Rates for Dividends and Capital Gains by Country

Country	Difference Between Top M.T.R.s for Dividends and Capital Gains
Australia	1.1
Canada	5.7
Chile	7.7
Denmark	0.0
Finland	(8.4)
France	4.3
Germany	1.4
Italy	(32.0)
Japan	0.0
Korea	31.1
Netherlands	25.0
New Zealand	6.9
Norway	0.0
Spain	(2.0)
Sweden	0.0
Switzerland	20.0
United Kingdom	8.1
United States	0.0

Note: Brackets indicate tax advantage for dividends vis-à-vis capital gains
Source: Carroll & Prante (2012), Ernst & Young LLP

6. MIGRATION RESPONSES TO CHANGES IN TAXATION

A) Introduction

Behavioural responses at the extensive margin may have an important impact on revenue collection, with common examples being migration and retirement.

Surprisingly, the impact of taxation on migration trends has not been studied in great detail. Studies which have been published are summarized in section 2.

Extensive margin responses such as migration are taken into account implicitly if the E.T.I. estimates are done via time series. This is because the time series models look at total taxable revenue declared in the country, and changes to the rates of the exit or entrance of individuals will impact this revenue measure. In order for these estimates to fully account for migration habits, we need the extra assumption that the average E.T.I. in the country is very nearly the same before and after any migration shift caused by a tax reform.

However, the situation is much more complex when we turn to panel data. Until now, no satisfactory model has been proposed which controls only for the component of migration or retirement planning which is a response to changes in taxation. If this component is at all significant relative to other behavioural responses, estimates of the E.T.I. and expected revenue loss will be biased, most likely downward.

To assess the likely magnitude of the problem, it is necessary to get a handle on exactly how responsive peoples' residency choices are to taxation. As previously mentioned, existing studies have found scant evidence of migration responses in general settings.

An excellent potential source of identification on this issue is the movement of labour in the Nordic countries, where cultural ties are strong.

Denmark, Finland, Iceland, Norway and Sweden have allowed free movement of labour among their countries since July 1st, 1954. There had already been bilateral agreements between countries prior to that date. Full rights in terms of social security were established by 1957, while free movement of labour in the health services industry was finalized in 1967. Pederson, Røed & Wadensjö (2008) provide a comprehensive summary of the evolution of these labour agreements.

Until the 1970s, most migration within Scandinavia was from Finland to Sweden. However, by that time the standard of living in Finland became more comparable to that of other Nordic countries, and the flow of migrants lessened.

Norway, which had been a net supplier of migrants in the 1960s and 1970s, had become a net receiver by the 1990s. This coincided with the expansion of the oil based economy.

Perhaps most advantageous in studying the Nordic countries is the high quality of available data.

Migration data, in terms of both immigration and emigration between each pair of countries, has been made available since 1966 in the Yearbook of Nordic Statistics. Additional pairwise migration data for Sweden from 1946 onwards was published in Wadensjö (1974). Information on marginal tax rates since 1956 comes from the O.E.C.D. tax database, as well as Statistics Norway and Du Rietz, Johansson & Stenkula (2013). Information on population levels and GDP growth since 1950 comes from the O.E.C.D. demographic and economic database, Edvinsson (2011) and Norges Bank Occasional Papers No. 35 (Historical Monetary Statistics for Norway 1819-2003).

A caveat is that for periods before 1970 there are discrepancies between data sources, particularly for calculations of all-in tax rates (those including social security contributions) and GDP growth. However, since these differences rarely exceed 5% in magnitude, it is unlikely that they invalidate the results, especially since the data covers an additional 43 years after that point, from 1970 to 2013.

Unfortunately, migration data is not divided by income levels. Hence it is not possible to examine the migration responses of top earners in isolation from the rest of the population. However, the cut-off to be in the top tax bracket is relatively low at less than $68,000. Additionally, those with such incomes have the means to immigrate to other countries, whereas lower income individuals may not have the financial means to move comfortably. Hence it is plausible that if there is a causal relationship between top tax rates and migration, a regression with the available data may be enough for identification.

I focus on migration between Sweden and Norway due to their extremely close cultural ties, and the availability of data from 1956 onwards. Results for migration between Denmark and Sweden produced similar results to that presented below, but I was only able to find data on marginal tax rates in Denmark from 1970 onwards. The smaller sample size produced results that were significantly less precise than those presented here, but which allowed similar inference.

The regression model is as follows:

$$\mathbf{mig}_t = \alpha + \textstyle\sum_i(\beta_i \Delta \mathbf{mtr}_{t-i}) + \delta_1 \Delta \mathbf{gdp}_{t-1} + \delta_2 t + \delta_3 t^2 + \varepsilon_t \quad i \in \{0, 1, 2, 3, 4, 5\} \tag{9}$$

Where:
mig_t is a measurement of migration in period t. It is defined as net migration, total immigration or total emigration from Sweden to Norway.
Δmtr_{t-i} is the difference in top marginal tax rates in period $t-i$ between Sweden and Norway.
Δgdp_{t-1} is the difference between GDP growth of Sweden and Norway in period t-1.
α is the regression constant
ε_t is the error term with a conditional expected value of zero in period t assuming correct specification of the model.

I control for differences in GDP growth instead of differences in unemployment rates since measures of GDP tend to be more consistent across countries. Incentives to return to the labour force vary greatly across countries, which impacts how many of those out of work are actively searching for a job and thus included in the unemployment rate. These incentives have also varied greatly within countries during the period studied (Carling et al (1996), Scarpetta (1996)).

Autocorrelation of the error terms is again a concern, so all regressions are presented with Newey-West standard errors, done with 3 lags.

Lastly, as with the work on capital gains, Almon distributed polynomial lags are used as a robustness check.

Figure 27 shows that visually there is no noticeable correlation between net migration rates and marginal tax rates:

Figure 27

Swedish Marginal Tax Rates and Net Nordic Migration

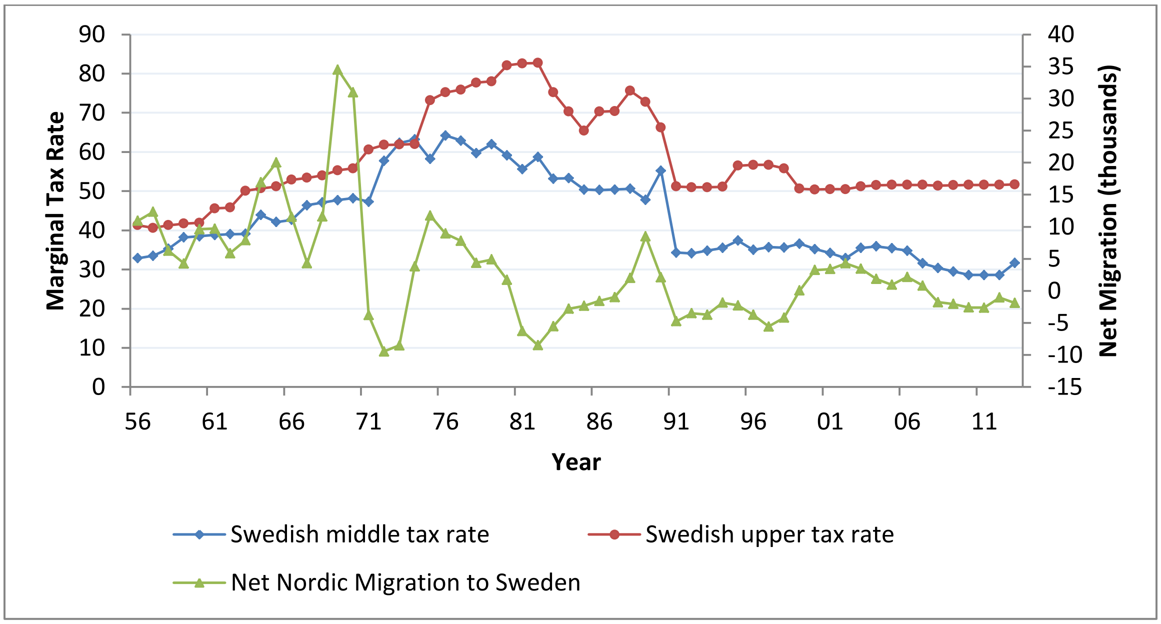


Figure 28 shows the same situation restricted to migration between Sweden and Norway:

Figure 28

Migration and Differences between Swedish and Norwegian Tax Rates

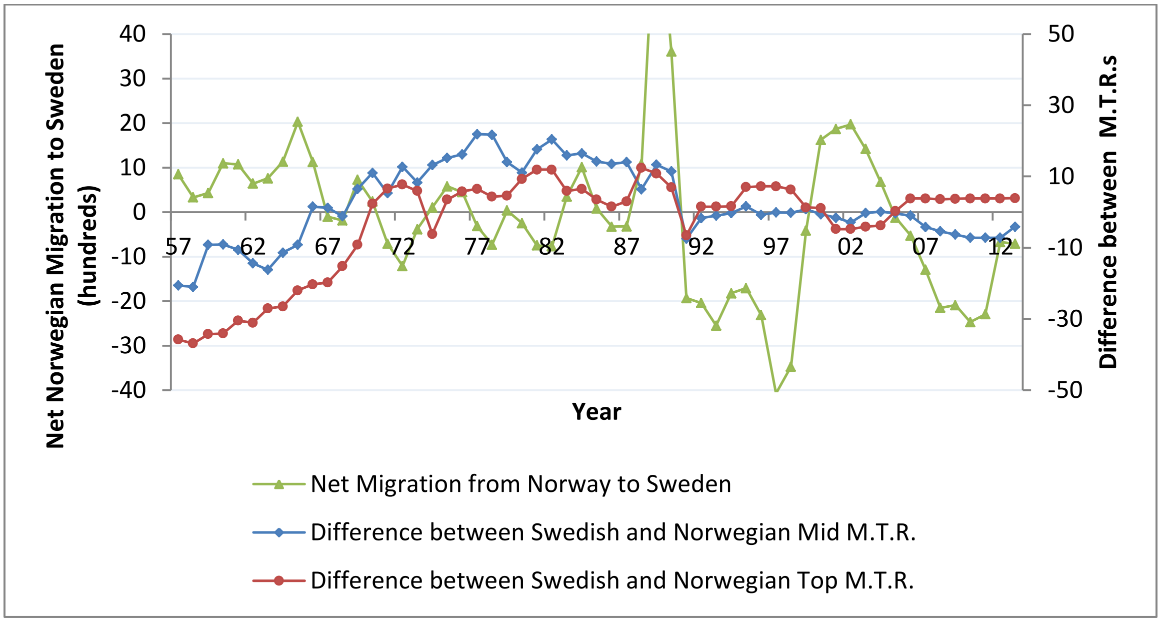
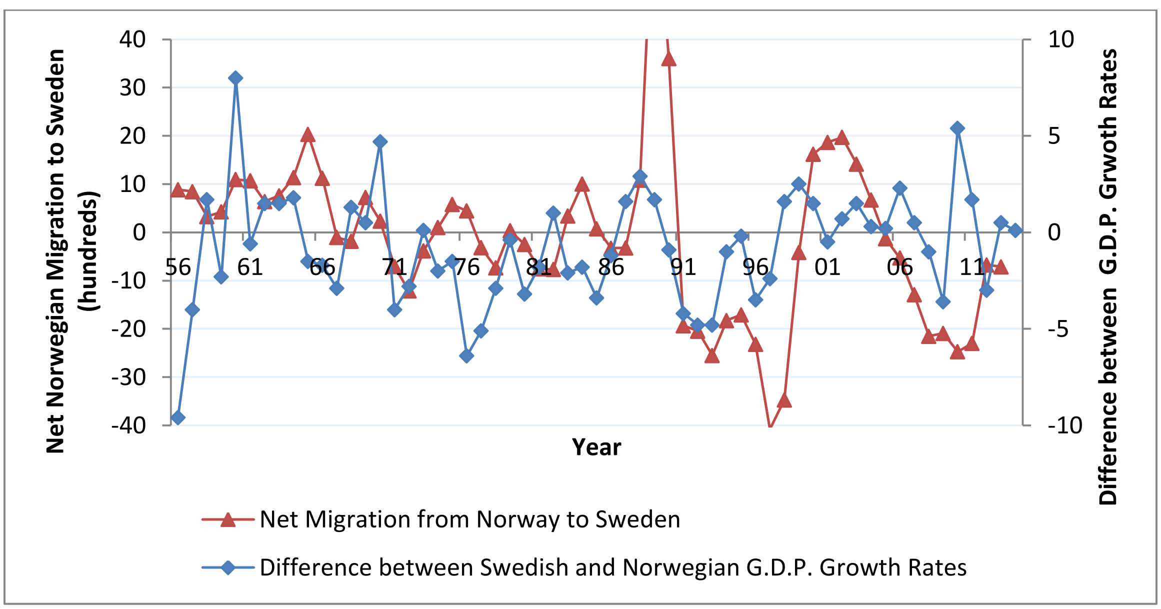


Figure 29 shows the relationship between GDP growth and net migration from Norway to Sweden. Here, a closer relationship is visually evident, providing justification for controlling for G.D.P. growth in the regression model.

Figure 29

Migration and Differences between Swedish and Norwegian GDP Growth Rates



Figures 30 and 31 show the composition of Nordic migration to and from Sweden. Whereas the Finnish used to account for a large share of the migrants, more recently the largest share has involved Norway.

Figure 30

Nordic Immigration to Sweden

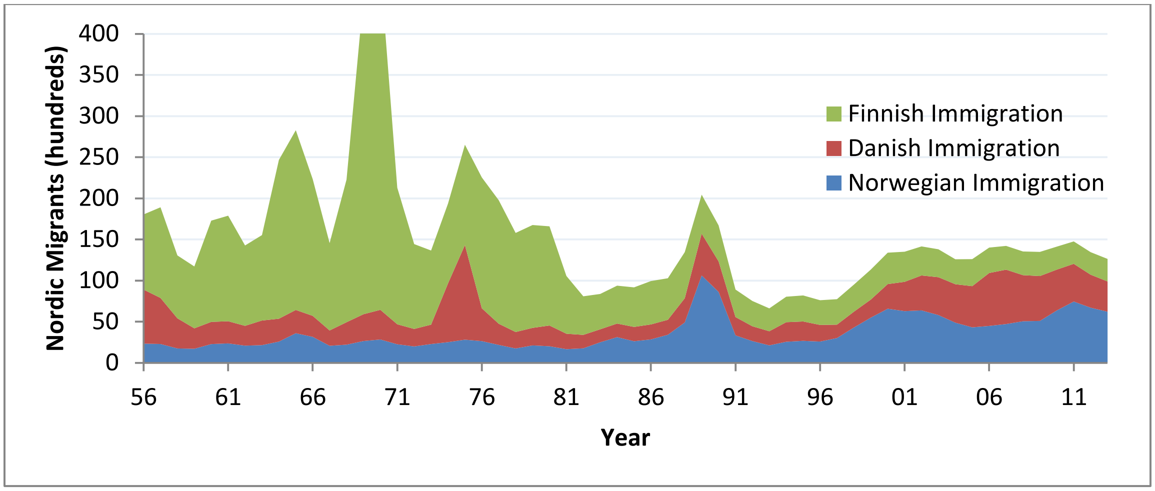
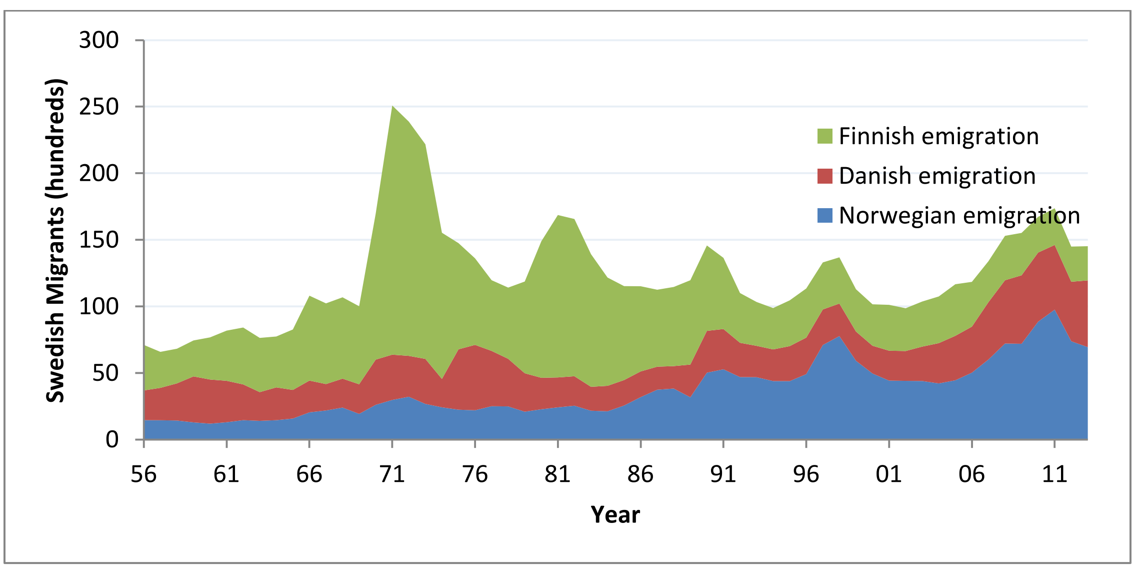


Figure 31

Nordic Emigration from Sweden



Note that data on migration does not specify citizenship of migrants, only which country they lived in and where they moved to. Hence the large exodus of residents of Sweden to Finland in the early 70s was likely driven by the Finns who moved to Sweden in the late 60s returning home.

These trends provide justification for controlling for time trends in the regression.

B) Results

Unless specifically stated, linear and quadratic time trend controls are significant at the 1% level in all regressions that follow.

References to top tax rates mean the top marginal rate in effect in each country.

For models involving net migration from Norway to Sweden, individual statistical significance appears only for lags of 4 and 5 years of the difference in marginal tax rates, with the signs of the estimates opposite to what we expect. That is, higher relative top tax rates in Sweden are predicted to increase immigration from Norway to Sweden after a period of 4 to 5 years.

This could be explained by the fact that migration rates do not distinguish between income classes. For instance, any exodus of high income individuals from Sweden could be masked by a greater influx of low to middle income Norwegians attracted by an expanding social security system.

In terms of polynomial methods and the net migration model, the large differences in estimates compared to a regular regression indicate that the polynomial technique has introduced large bias, which is usually because the polynomials do not fit the data well. Focusing exclusively on the standard errors, the polynomial technique of degree 2 is somewhat successful at increasing precision, whereas the degree 3 polynomials do not even provide significant improvements in this sense.

Table 22

Regression in Levels of Net Migration from Norway to Sweden on the Difference in Top Tax Rates and Almon Distributed Polynomial Lag Technique

1956 to 2013

Years of Lag of the Difference in Tax Rates	Regular Regression with linear time trend control	Regular Regression with quadratic time trend control	Degree 2 Polynomial with linear time trend control	Degree 2 Polynomial with quadratic time trend control	Degree 3 Polynomial with linear time trend control	Degree 3 Polynomial with quadratic time trend control
0	-52.07 (30.84)	-51.78 (31.80)	-10.18 (43.26)	-10.38 (41.66)	-47.33 (30.45)	-47.19 (30.56)
1	-0.29 (68.87)	-0.0056 (68.26)	-26.16 (13.25)	-26.25 (13.15)	8.24 (31.03)	8.86 (29.26)
2	41.25 (45.12)	42.01 (48.17)	-25.70 (16.26)	-25.76 (17.37)	-4.71 (13.18)	-4.30 (13.71)
3	-99.98 (70.59)	-99.82 (71.92)	-8.79 (18.34)	-8.93 (20.98)	-29.72 (33.07)	-29.55 (34.25)
4	42.45 (28.10)	42.77 (28.54)	24.56 (11.02)	24.25 (17.40)	-10.28 (31.00)	-9.80 (34.30)
5	95.16 (39.93)	97.38 (40.99)	74.36 (29.69)	73.78 (31.63)	110.08 (52.44)	112.07 (50.02)

Table 23 presents the sum of the effects from each year of lag, which can be seen as the total effect on net migration of an increase in the relative taxation of high income individuals in Sweden. The units of measurement are individual migrants. All results show an increase in the net flow of migrants from Norway to Sweden, although of a very small magnitude. The same table presents 95% confidence intervals, which show that the probabilities of large-magnitude changes in net migration rates due to modest to intermediate changes in tax policy are small.

Table 23

Responsiveness of Net Migration from Norway to Sweden to the Difference in Top Tax Rates
Period of 5 Years
(Measured in Number of Migrants with a 95% Confidence Interval)

	Regular Regression with linear time trend control	Regular Regression with quadratic time trend control	Degree 2 Polynomial with linear time trend control	Degree 2 Polynomial with quadratic time trend control	Degree 3 Polynomial with linear time trend control	Degree 3 Polynomial with quadratic time trend control
5 year effect	26.5	30.6	28.1	26.7	26.3	30.1
S. D.	31.6	50.5	33.9	48.6	32.3	49.2
Conf. Int.	(-35.4, 88.4)	(-68.4, 129.5)	(-38.4, 94.6)	(-68.5, 122.0)	(-37.1, 89.6)	(-66.4, 126.6)

Looking at total migration from Norway to Sweden, we see little evidence of any impact from changes in relative taxation over 5 year periods. Signs of coefficients tend to switch unexpectedly, but very few of the estimates are significant at the 10% level, and none are significant at the 5% level.

The polynomial technique applied to this model has similar issues to that of the net migration model. None of the polynomial techniques substantially increase precision. Yet the degree 3 polynomials tend to be more comparable with the regular regressions in terms of the magnitudes of the estimates of coefficients, whereas the degree 2 polynomials often lose almost all precision.

Table 24

Regression in Levels of Total Migration from Norway to Sweden on the Difference in Top Tax Rates and Almon Distributed Polynomial Lag Technique

1956 to 2013

Years of Lag of the Difference in Tax Rates	Regular Regression with linear time trend control	Regular Regression with quadratic time trend control	Degree 2 Polynomial with linear time trend control	Degree 2 Polynomial with quadratic time trend control	Degree 3 Polynomial with linear time trend control	Degree 3 Polynomial with quadratic time trend control
0	-46.74 (28.41)	-44.39 (28.25)	-6.03 (36.83)	-2.19 (33.08)	-41.89 (25.21)	-40.71 (24.39)
1	19.37 (59.70)	21.68 (58.54)	-6.42 (9.07)	-4.76 (8.35)	26.79 (34.48)	31.97 (32.32)
2	58.96 (40.72)	65.03 (40.08)	-4.71 (12.41)	-3.41 (12.98)	15.54 (14.7)	19.05 (14.62)
3	-88.49 (67.60)	-87.16 (68.69)	-0.91 (15.3)	1.85 (17.57)	-21.11 (31.03)	-19.73 (31.98)
4	22.32 (25.37)	24.88 (24.51)	4.97 (8.29)	11.02 (13.66)	-28.65 (32.16)	-24.61 (35.59)
5	32.88 (41.67)	50.68 (31.69)	12.95 (22.1)	24.11 (16.69)	47.43 (50.66)	64.17 (40.9)

In terms of total effects over the 5 year period, we see once more through the confidence intervals that the total magnitude of impacts on immigration rates of Norwegians to Sweden is almost certainly low. This implies that rates of taxation do not figure prominently among factors affecting decisions to move from Norway to Sweden.

Table 25

Responsiveness of Immigration from Norway to Sweden to the Difference in Top Tax Rates
Period of 5 Years
(Measured in Number of Migrants with a 95% Confidence Interval)

	Regular Regression with linear time trend control	Regular Regression with quadratic time trend control	Degree 2 Polynomial with linear time trend control	Degree 2 Polynomial with quadratic time trend control	Degree 3 Polynomial with linear time trend control	Degree 3 Polynomial with quadratic time trend control
5 year effect	-1.7	30.7	-0.2	26.6	-1.9	30.2
S. D.	35.9	36.5	37.6	35.2	36.1	36.4
Conf. Int.	(-72.0, 68.6)	(-40.9, 102.3)	(-73.9, 73.6)	(-42.3, 95.6)	(-72.7, 68.9)	(-41.1, 101.4)

Switching to emigration from Sweden to Norway, we again notice little impact of a substantial effect of rates of taxation. Magnitudes of coefficient estimates from regular regressions are all low, and the only ones which are statistically significant at conventional levels are those of the 5 year lag term, which are of the opposite sign to what we expect.

We do however see a noticeable improvement of the estimates of the polynomial technique for the model analysing migration from Sweden to Norway. The estimates are similar to those via a regular regression, meaning the bias introduced by imposing the polynomial relationship on the data appears to be much less than in the previous two models. Additionally, the t-statistics are larger in absolute value (though not necessarily significant at even the 10% level), meaning that the polynomial technique has succeeded in at least improving precision. The coefficients are positive for lower lags, but only those of a degree 2 polynomial technique for the coefficient of a 1 year lag are significant at even the 5% level. They mostly become negative as the lag length increases. Still, the magnitude of the coefficients in absolute value is small in all cases.

Table 26

Regression in Levels of Total Migration from Sweden to Norway on the Difference in Top Tax Rates and Almon Distributed Polynomial Lag Technique

1956 to 2013

Years of Lag of the Difference in Tax Rates	Regular Regression with linear time trend control	Regular Regression with quadratic time trend control	Degree 2 Polynomial with linear time trend control	Degree 2 Polynomial with quadratic time trend control	Degree 3 Polynomial with linear time trend control	Degree 3 Polynomial with quadratic time trend control
0	5.33 (20.57)	7.39 (21.65)	4.15 (14.03)	8.19 (15.77)	5.44 (20.3)	6.48 (21.17)
1	19.67 (20.45)	21.69 (19.64)	19.74 (7.97)	21.49 (8.34)	18.55 (13.59)	23.12 (13.31)
2	17.71 (27.83)	23.02 (29.52)	20.99 (11.94)	22.35 (12.52)	20.26 (14.72)	23.35 (15.33)
3	11.49 (20.65)	12.66 (21.16)	7.87 (11.21)	10.78 (12.38)	8.6 (11.88)	9.83 (12.15)
4	-20.13 (22.69)	-17.89 (23.76)	-19.59 (6.85)	-13.23 (11.14)	-18.38 (12.21)	-14.81 (14.15)
5	-62.28 (20.52)	-46.70 (24.03)	-61.41 (16.83)	-49.68 (23.61)	-62.65 (18.9)	47.9 (23.86)

Table 27 shows that total effects on migration from Sweden to Norway over the 5 year period are predicted to be quite modest, and of the counter-intuitive sign. This is especially surprising, as increased immigration of middle to low income Norwegians cannot be an explanation for low responsiveness in the case of total emigration from Sweden to Norway following a top tax rate increase in Sweden.

One possibility is that improved social security programs in Sweden reduce the numbers of middle class Swedes moving to Norway.

Perhaps a more credible explanation is that high income Swedes may move their assets out of Sweden while remaining residents of the country. This can be either through legal means (transferring income to non-resident trusts or moving a corporation's activities to a lower-tax district) or illegally hiding income offshore in tax havens. Such financial strategies could significantly reduce the sensitivity of residency decisions to taxation policies, while still resulting in large elasticities of taxable income for high earners.

Table 27

Responsiveness of Emigration from Sweden to Norway to the Difference in Top Tax Rates
Period of 5 Years
(Measured in Number of Migrants with a 95% Confidence Interval)

	Regular Regression with linear time trend control	Regular Regression with quadratic time trend control	Degree 2 Polynomial with linear time trend control	Degree 2 Polynomial with quadratic time trend control	Degree 3 Polynomial with linear time trend control	Degree 3 Polynomial with quadratic time trend control
5 year effect	-28.2	0.2	-28.2	-0.1	-28.2	0.1
S. D.	15.2	41.9	14.6	40.3	14.8	40.7
Conf. Int.	(-57.9, 1.5)	(-81.9, 82.3)	(-56.8, 0.3)	(-79.1, 78.9)	(-57.1, 0.7)	(-79.8, 79.9)

Table 28 highlights the difficulty in rejecting the hypothesis that there is no response at all of migration decisions to changes in top tax rates when using traditional regression methods. The only scenario where we can reject the null hypothesis at the 5% level is with Swedish emigration and a linear time control variable included in the regression, but even then, the estimate of the coefficient is of the opposite sign to what we would expect.

Table 28

Responsiveness of Migration from Sweden to Norway to the Difference in Top Tax Rates
P-Values for Null Hypothesis: No Impact During a 5 Year Period

	Linear Time Trend Control	Quadratic Time Trend Control
Net Migration	0.0772	0.1648
Immigration	0.3595	0.1943
Emigration	0.0047	0.2543

We therefore have no new evidence to suggest a bias due to the exclusion of migration in estimates of the E.T.I. done with panel data. This is broadly in line with the results on this topic published in the past, as summarized in section 2.

Even taking a value such as a drop in net migration of 80 people, which is at the high end of the 95% confidence intervals which have been constructed, we do not see appreciable amounts of lost revenue. Assuming a marginal tax rate of 60% and $250,000 of annual revenue taxed at the top marginal tax rate, a country such as Sweden would only lose $12,000,000 in revenue each year due to emigration. This is the equivalent of less than $2 per citizen and far less than the extra income generated by the tax increase.

If evidence is not forthcoming in the case of Norway and Sweden, two countries with very strong social, cultural and economic ties which publish an abundance of data, it is difficult to imagine it being easier to find in other geographical regions using similar forms of data.

However, in a scenario where migration data were differentiated by income class, results of similar regressions as those presented here could be much more informative.

In the absence of such detailed data, a more effective approach to studying responses at the extensive margin and the flight of capital following tax increases would be to look at various sheltering techniques readily available to high income individuals. Such studies may stand a much better chance of quantifying the extensive margin responses of taxable income to changes in marginal tax rates.

7. CONCLUSIONS AND AREAS FOR FURTHER RESEARCH

We have seen how important an understanding of behavioral responses is in policy planning. Yet accurately estimating the E.T.I. which describes such responses is quite difficult. Even a change in the elasticity by half a point has drastic impacts on revenue collection. Large amounts will be lost to behavioral responses as people reduce their taxable income following a tax hike.

However, it is becoming increasingly clear that most reactions of taxable income following a change in marginal tax rates occur through the tax base. This is reflected in the results of this thesis on capital gains and migration. Taxpayers are quick to rearrange their investment portfolios to take advantage of preferential tax rates, but no evidence has been found showing a tendency to physically relocate their residence due to tax considerations.

Chetty et al (2011) states that 60% of Danish taxpayers have no tax deductions at all. Furthermore, as illustrated in section 3, studies have consistently found very low estimates of the E.T.I. in Denmark. The situation in other Scandinavian countries is similar. Additionally, marginal tax rates in these regions are quite high, and kick in at relatively low income levels.

Hence by broadening the tax base as much as possible, the E.T.I. will almost certainly fall. Revenue forecasting would then be much easier and far more accurate, since the E.T.I. is so difficult to estimate correctly. Accompanying such reforms by cuts to marginal tax rates could make the changes revenue neutral.

It would be very useful to develop regression models which can better control for the tax base. Having a detailed comparison of measures of the tax base across the countries studied in this thesis would also be of great interest.

It would also be quite informative to examine how changes to average tax rates affect behavioral responses, as such rates are much harder to calculate. Such a model would necessarily have to incorporate payroll and sales taxes, which are substitutes for individual income taxes in terms of revenue collection for the government. Yet all contribute to how much disposable income a citizen truly has.

Lastly, migration data differentiated by individual income would allow a much more precise study of migration habits of the wealthy following tax changes. Such results would not be confounded by possibly opposing migration of other income classes.

It is certain that with the current budgetary difficulties of governments across the globe, further research on this topic is more pertinent than ever.

REFERENCES

Aarbu, Karl O., and Thor O. Thoresen. "Income responses to tax changes—evidence from the Norwegian tax reform." National Tax Journal (2001): 319-335.

Almon, Shirley, "The distributed lag between capital appropriations and net expenditures, "Econometrica 33, 1965, 178 – 196

Arrazola, María, et al. "Personal Income Tax Reforms and the Elasticity of Reported Income to Marginal Tax Rates: An Empirical Analysis Applied to Spain." (2014).

Auerbach, Alan J. 1988. "Capital Gains Taxation in the United States: Realizations, Revenue, and Rhetoric." Brookings Papers on Economic Activity 2: 595–631.

Auerbach, Alan J., Leonard E. Burman, and Jonathan Siegel. "Capital gains taxation and tax avoidance: New evidence from panel data." No. w6399. National Bureau of Economic Research, 1998.

Auten, Gerald and Robert Carroll. 1999 "The Effect of Income Taxes on Household Behavior." Review of Economics and Statistics, 81(4): 681-693.

Bach, Stefan, Giacomo Corneo, and Viktor Steiner. Top incomes and top taxes in Germany. No. 1641. CESifo working papers, 2006.

Bahl, Roy, Jorge Martinez-Vazquez, and Sally Wallace. "State and local government choices in fiscal redistribution." National Tax Journal (2002): 723-742.

Bakija, Jon, and Joel Slemrod. "Do the rich flee from high state taxes? Evidence from federal estate tax returns". No. w10645. National Bureau of Economic Research, 2004.

Bastani, Spencer, and Håkan Selin. "Estimating the taxable income elasticity at kink points of the Swedish tax system." (2011).

Brewer, Mike, Emmanuel Saez, and Andrew Shephard. "Means-testing and tax rates on earnings, in 'The Mirrlees Review: Reforming the Tax System for the 21st Century'." (2008).

Bruce, Donald, William F. Fox, and Zhou Yang. "Base mobility and state personal income taxes." National Tax Journal 63.4 (2010): 945-966.

Cabannes, Pierre-Yves, Cedric Houdre, and Camille Landais. "How does the taxable income of households react to taxation? An estimate for the period 1997-2004." Economie & Statistique (2014).

Campano, Fred, and Dominick Salvatore. Income Distribution: Includes CD. Oxford University Press, 2006.

Carey, Simon, et al. "Estimating the Elasticity of Taxable Income in New Zealand." Economic Record 91.292 (2015): 54-78.

Carling, Kenneth, et al. "Unemployment duration, unemployment benefits, and labor market programs in Sweden." Journal of Public Economics 59.3 (1996): 313-334.

Carroll, Robert, and Gerald Prante. "Corporate Dividend and Capital Gains Taxation: A comparison of the United States to other developed nations." Ernst & Young (February) http://images.politico.com/global/2012/02/120208_asidividend.html (2012).

Champernowne, David G. "The graduation of income distributions."Econometrica: Journal of the Econometric Society (1952): 591-615.

Chetty, Raj, et al. "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records." The quarterly journal of economics 126.2 (2011): 749.

Chetty, Raj. "Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply." Econometrica 80.3 (2012): 969-1018.

Chetty, Raj. 2009a. "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply." National Bureau of Economic Research Working Paper 15616.

Chetty, Raj. 2009b. "Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implications of Evasion and Avoidance." American Economic Journal: Economic Policy 1 (2): 31–52.

Chetty, Raj. 2009c. "Sufficient Statistics for Welfare Analysis: A Bridge between Structural and Reduced-Form Methods." Annual Review of Economics 1 (1): 451–87.

Conway, Karen Smith and Andrew J. Houtenville "Do the Elderly "Vote with Their Feet?" Public Choice, Vol. 97, No. 4 (1998), pp. 663-685

Conway, Karen Smith, and Andrew J. Houtenville. "Elderly migration and state fiscal policy: evidence from the 1990 census migration flows." National Tax Journal (2001): 103-123.

Conway, Karen Smith, and Jonathan C. Rork. State "Death" Taxes and Elderly Migration–The Chicken or the Egg?. Working paper, 2005.

Coomes, Paul A., and William H. Hoyt. "Income taxes and the destination of movers to multistate MSAs." Journal of Urban Economics 63.3 (2008): 920-937.

Dagum, C., A New Model of Personal Income Distribution: Specification and Estimation, Economie appliquée, Vol. 30, 1977, pp. 413-437

Day, Kathleen M., and Stanley L. Winer. "Policy-induced internal migration: An empirical investigation of the Canadian case." International Tax and Public Finance 13.5 (2006): 535-564.

Della Croce, R., Stewart, F., & Yermo, J. (2011). Promoting Longer-Term Investment by Institutional Investors. OECD Journal: Financial Market Trends,2011(1), 145-164.

Dixon, Sylvia. "The distribution of earnings in New Zealand." Labour Market Bulletin 1 (1996): 45-100.

Doerrenberg, Philipp, Andreas Peichl, and Sebastian Siegloch. "Sucient statistic or not? The elasticity of taxable income in the presence of deduction possibilities." The Elasticity of Taxable Income in the Presence of Deduction Possibilities (October 7, 2014). ZEW-Centre for European Economic Research Discussion Paper 14-078 (2014).

Du Rietz, G., Johansson, D. and Stenkula, M., 1862. Swedish labor income taxation (1862–2013). Swedish Taxation: Developments Since.

Edvinsson, R., 2011. New estimates of Swedish GDP by activity 1665-2010. Stockholm Papers in Economic History, (12).

Feldstein, Martin and Marian Vaillant. 1994. "Can state taxes redistribute income?" Journal of Public Economics 68(3): 369-396.

Feldstein, Martin. 1995. "The 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.

Fernandez, Camila Nino. "The elasticity of taxable income: The cases of Chile and Colombia."

Gagné, Robert, François Vaillancourt, and Jean-François Nadeau. "Taxpayers' response to tax rate changes: a Canadian panel study." CIRANO, 2000.

Giertz, Seth H. "The elasticity of taxable income: influences on economic efficiency and tax revenues, and implications for tax policy." Economics Department Faculty Publications (2009): 64.

Gottfried, Peter, and Daniela Witczak. "The responses of taxable income induced by tax cuts: empirical evidence from the German taxpayer panel." No. 57. IAW-Diskussionspapiere, 2012

Graham, Benjamin. "The Intelligent Investor" (2005)

Gruber, Jon, and Emmanuel Saez. 2002. "The Elasticity of Taxable Income: Evidence and Implications." Journal of Public Economics 84 (1): 1–32.

Haldane, A. (2010, September). Patience and finance. In Remarks at the Oxford China Business Forum, Beijing, available at
http://www. bankofengland. co. uk/publications/speeches/2010/speech445.pdf.

Hansson, Åsa. "Taxpayers' responsiveness to tax rate changes and implications for the cost of taxation in Sweden." International Tax and Public Finance 14.5 (2004): 563-582.

Heim, Bradley T. 2009. "The Effect of Recent Tax Changes on Taxable Income: Evidence from a New Panel of Tax Returns." Journal of Policy Analysis and Management 28 (1): 147–63.

Holmlund, Bertil, and Martin Söderström. "Estimating dynamic income responses to tax reform." The BE Journal of Economic Analysis & Policy 11.1 (2011).

Holmlund, Bertil, and Martin Söderström. "Estimating dynamic income responses to tax reform." The BE Journal of Economic Analysis & Policy 11.1 (2011).

Hourton, Alain. Income Inequality in Chile: 1990-2006. No. e-12-004. 2012.

Hunter, G., M. Sanchez, and F. Douglas. "Incomes of the one percent and everyone else in Canada." Social Policy Research Unit (2012).

Imbens, G. W., and T. Lancaster, ''Efficient Estimation and Stratified Sampling,'' Journal of Econometrics 74 (1996), 289–318.

Kirchgässner, Gebhard, and Werner W. Pommerehne. "Tax harmonization and tax competition in the European Union: Lessons from Switzerland." Journal of Public Economics 60.3 (1996): 351-371.

Kleintop, J. (2012). What the Fed Can't Do. LPL Financial Research, Weekly Market Commentary, 6.

Kleven, Henrik Jacobsen, et al. "Migration and wage effects of taxing top earners: evidence from the foreigners' tax scheme in Denmark." No. w18885. National Bureau of Economic Research, 2013.

Kleven, Henrik, and Esben Anton Schultz. 2011. "Estimating Taxable Income Responses using Danish Tax Reforms."
http://personal.lse.ac.uk/kleven/Downloads/MyPapers/WorkingPapers/klevenschultz_jan2012.pdf.

Klovland, J.T. and Qvigstad, J.F., 2004. Historical Monetary Statistics for Norway 1819-2003 (Vol. 35, pp. 241-88). Oslo: Norges bank.

Knapp, Thomas A., Nancy E. White, and David E. Clark. "A nested logit approach to household mobility." Journal of Regional Science 41.1 (2001): 1-22.

Kopczuk, Wojciech. 2005. "Tax Bases, Tax Rates and the Elasticity of Reported Income." Journal of Public Economics, 89(11-12): 2093-2119.

Le Maire, Daniel, and Bertel Schjerning. "Tax bunching, income shifting and self-employment." Journal of Public Economics 107 (2013): 1-18.

Lindsey, Lawrence. 1987. "Individual Taxpayer Response to Tax Cuts: 1982-1984, with Implications for the Revenue Maximizing Tax Rate." Journal of Public Economics, 33: 173-206.

Martínez, Isabel Z. "Should I Stay or Should I Go? Tax-Induced Mobility and the Taxable Income Elasticity in Switzerland." (2014).

Massarrat-Mashhadi, N., and C. Werdt. "Estimating dynamic income responses to tax changes Massarrat-Mashhadi: Evidence from Germany Free University Berlin Working Paper." (2012).

Matikka, Tuomas. "Taxable income elasticity and the anatomy of behavioral response: Evidence from Finland." Government Institute for Economic Research Working Papers 55 (2014).

Mendoza, Manuel Díaz. "La respuesta de los contribuyentes ante las reformas del IRPF ", 1987-1994. CEMFI, 2004.

Milligan, Kevin, and Michael Smart. "Estimating high-income tax elasticities using sub-national variation in tax rates." (2014).

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

Milligan, Kevin, and Michael Smart. "Taxation and top incomes in Canada." No. w20489. National Bureau of Economic Research, 2014.

Newey, Whitney K., and Kenneth D. West. "Hypothesis testing with efficient method of moments estimation." International Economic Review (1987): 777-787

Pedersen, P.J., Roed, M. and Wadensjö, E., 2008. The common Nordic labour market at 50. Nordic Council of Ministers.

Rogerson, Richard, and Johanna Wallenius. "Micro and macro elasticities in a life cycle model with taxes." Journal of Economic theory 144.6 (2009): 2277-2292.

Saez, Emmanuel, Joel Slemrod, and Seth Giertz, ''The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review,'' Journal of Economic Literature, 50 (2012), 3–50.

Saez, Emmanuel. 2004b. "Reported Incomes and Marginal Tax Rates, 1960–2000: Evidence and Policy Implications." In Tax Policy and the Economy, Volume 18, edited by James M. Poterba, 117–74. Cambridge and London: MIT Press.

Saez, Emmanuel. 2010. "Do Taxpayers Bunch at Kink Points?" American Economic Journal: Economic Policy 2 (3): 180–212.

Sanandaji, Tino (2012) : "The International Mobility of the Super-Rich", IFN Working Paper, No. 904

Sanz-Sanz, José Félix, et al. "Reported gross income and marginal tax rates: estimation of the behavioural reactions of Spanish taxpayers." Applied Economics 47.5 (2015): 466-484.

Scarpetta, Stefano. "Assessing the role of labour market policies and institutional settings on unemployment: A cross-country study." OECD Economic studies 26.1 (1996): 43-98.

Selén, Jan. "Taxable Income Responses to Tax Changes-A Panel Analysis of the 1990/91 Swedish Reform." No. 177. 2002.

Sillamaa, Mary-Anne Sillamaa Anne, and Michael R. Veall Michael R. Veall. "Social And Economic Dimensions Of An Aging Population." (2000).

Slemrod, Joel and Wojciech Kopczuk. 2002. "The Optimal Elasticity of Taxable Income." Journal of Public Economics, 84(1): 91-112.

Slemrod, Joel. 1996. "High Income Families and the Tax Changes of the 1980s: The Anatomy of Behavioral Response." In Empirical Foundations of Household Taxation, eds. Martin Feldstein and James Poterba, University of Chicago.

Smart, Michael. "Taxation and deadweight loss in a system of intergovernmental transfers." Canadian Journal of Economics (1998): 189-206.Solon, Gary, Steven J. Haider, and Jeffrey M. Wooldridge. "What are we weighting for?." Journal of Human Resources 50.2 (2015): 301-316.

Sousa-Poza, Alfonso, Patrick A. Puhani, and Thomas Liebig "Taxation and Internal Migration-Evidence from the Swiss Census using Community-Level Variation in Income Tax Rates." No. 348. Diskussionspapiere des Fachbereichs Wirtschaftswissenschaften, Universität Hannover, 2006.

Thomas, Alastair. "Taxable income elasticities and the deadweight cost of taxation in New Zealand." New Zealand Association of Economists Annual Conference, Christchurch, June. 2007.

Thoresen, Thor O. "Reduced tax progressivity in Norway in the nineties: The effect from tax changes." International Tax and Public Finance 11.4 (2002): 487-506.

Thoresen, Thor O., and Trine E. Vattø. "Validation of structural labor supply model by the elasticity of taxable income." No. 738. 2013.

Vattø, Trine Engh. "Labor Earnings Responses to Income Taxation Empirical Evidence from Norway." (2012).

Wadensjö, Eskil (1974), "Immigrationen" in SOU 1974:29, Att utvärdera arbetsmarknadspolitik

Weber, Caroline E. "Toward obtaining a consistent estimate of the elasticity of taxable income using difference-in-differences." Journal of Public Economics 117 (2014): 90-103.

Young, Cristobal, and Charles Varner. "Millionaire migration and state taxation of top incomes: Evidence from a natural experiment." National Tax Journal 64.2 (2011): 255-284.

Zellner, Arnold. "An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias." Journal of the American statistical Association 57.298 (1962): 348-368.

APPENDIX

A) Technical Specifications of Estimating the E.T.I.

What follows are explicit statements of the three most popular regression models for obtaining estimates of the E.T.I. Unless otherwise stated, the articles do not include discussions of the properties of the error terms of the regressions.

Also, notation is listed so as to match that of the original articles.

The first model is that which was proposed for panel data by Auten & Carroll (1999), who start with:

$$\log(Y_{it}) = 1_i + \gamma_t + \alpha_t X_i + \beta \log(1\text{-}\tau_{it}) + \text{error}_{it} \tag{A1}$$

Where:
Y_{it} is the income of individual i in period t
1_i is an individual-specific effect
γ_t is a time-specific effect
X_i is a set of individual characteristics that do not change over time, but whose relationship to income may change over time. This is reflected by the α_t coefficients.
$1\text{-}\tau_{it}$ is the net-of-tax rate for individual i in period t.
The error term has a conditional expected value of zero assuming the rest of the model is correctly specified.

Taking first differences of (A1) yields:

$$\Delta\log(Y_i) = \Delta\gamma_t + \Delta\alpha X_i + \beta\Delta\log(1\text{-}\tau_i) + \varepsilon_{it} \tag{A2}$$

where Δ denotes the change in a variable between period t and t–1, and the error term ε_{it} has a conditional expected value of zero.

The tax rate $\boldsymbol{\tau_i}$ is estimated using the TAXSIM calculator, maintained by the National Bureau of Economic Research. This is the standard tax rate calculator when working with data from the U.S.

As stated above, the instrument for the change in the net-of-tax rate is what the net-of-tax rate would have been in period t had income in period t-1 only increased by the amount of inflation. Lastly, a taxpayer's income in period t-1 is usually included as a control variable to account for mean reversion and broad income trends.

The estimate of β in (A2) is the E.T.I.

The second standard model for panel data is a slight variation on the first, proposed by Gruber & Saez (2002):

$$\log(z_{i2}/z_{i1}) = \alpha_0 + \vartheta \log[(1-T'_{i2}(z_{i2}))/(1-T'_{i1}(z_{i1}))]$$
$$+ \eta \log [(z_{i2} - T_{i2}(z_{i2}))/ (z_{i1} - T_{i1}(z_{i1}))] + \alpha_1 \log(z_{i1}) + \Sigma_k \alpha_{2ik} mars_{ik}$$
$$+ \Sigma_j \alpha_{3j} YEAR_j + \Sigma_l \alpha_{4l} SPLINE_l(z_{i1}) + \varepsilon_{it} \qquad \text{(A3)}$$

Where:
z_{it} is before tax income in period t for individual i.
$T'_{it}(z_{it})$ is the marginal tax rate in period t for individual i.
$T_{it}(z_{it})$ is the total income tax paid in period t for individual i.
$z_{it} - T_{it}(z_{it})$ is the after tax income in period t for individual i.
$mars_{ik}$ is a set of marital status dummies in the base year.
$YEAR_j$ denote base year dummies.
$SPLINE_l(z_{i1})$ denotes a 10-piece spine in log first period income, where the l-th spline varies along with log income inside of the l-th income decile, and is constant outside that interval.
ϑ is the compensated elasticity of taxable income (which therefore measures the substitution effect).
η is the income effect parameter.
ε_{it} is the error term, mean zero under correct specification assumptions.

In particular, the spline is an effort to control for income trends and mean reversion.

The equation is then estimated by 2SLS using $\log[(1-T'_{ip}(z_{i2}))/(1-T'_{it}(z_{i1}))]$ and $\log [(z_{i1} - T_{i2}(z_{i1}))/ (z_{i1} - T_{i1}(z_{i1}))]$ as instruments.

Here, T'_p is the marginal tax rate in year t faced by the taxpayer if his income did not change in real terms from period t-1. The indexing using p is a convention introduced by the authors. Hence this instrument is the same concept as that used in equation (A2).

In practice it has been shown that η is not statistically different from zero, so equation (A3) is often simplified to

$$\log(z_{i2}/z_{i1}) = \alpha_0 + \vartheta \log[(1-T'_{i2}(z_{i2}))/(1-T'_{i1}(z_{i1}))] + \alpha_1 \log(z_{i1})$$
$$+ \Sigma_k \alpha_{2ik} mars_{ik} + \Sigma_j \alpha_{3j} YEAR_j + \Sigma_l \alpha_{4l} SPLINE_l(z_{i1}) + \varepsilon_{it} \qquad \text{(A4)}$$

As is shown later in this section, studies emulating the techniques of Gruber & Saez (2002) have used either equation (A3) or (A4), depending on whether the purpose of the study was to examine income effects along with substitution effects. Studies are categorized accordingly in table 3.

Since there are many controls, it is necessary to have access to a set of panel data containing multiple years in order to avoid destruction of identification. Gruber & Saez (2002) use stacked observations from nine pairs of years (11 consecutive years of data and three years difference in the pair, so that *t* and *t*-1 are three years apart), and present estimates that correct the standard errors in case of individual-specific correlation in how incomes change over time. Also, the authors weight all estimates by income to reflect the relative contribution to total revenues.

Their estimate for the E.T.I. ϑ is 0.40, as opposed to the 0.55 found by Auten & Carroll. They also find an estimate of the elasticity of broad income with respect to the net-of-tax rate of 0.12.

Finally, the third commonly used regression technique comes from the aforementioned Saez (2004) and simple time-series regressions involving time trend controls. Specifically, Saez relies on aggregate data released annually by the I.R.S. since 1960 in place of traditional panel data. The log of either income or income shares is regressed on the log of the net-of-tax rate along with time trends (linear, quadratic and cubic). Standard errors are Newey-West with 8 lags. Endogeneity of the log of net-of-tax rate is overcome via the log of one minus the top marginal income tax rate as an instrument. He uses as income A.G.I. (adjusted gross income, which is total gross income less specific deductions) minus capital gains, as opposed to taxable income. Taxable income is therefore adjusted gross income less allowances for personal exemptions and itemized deductions. Saez's estimate for the E.T.I. for the American population as a whole is a statistically insignificant 0.2. For the top 1% of the income distribution he estimates a statistically significant E.T.I. value of 0.5.

The results of Saez (2004) vary substantially by decade. In periods such as the 1960s when the tax base was broader, the E.T.I. estimates are low. Similarly, during periods such as the late 1970s and 1980s when more deductions were available, E.T.I. estimates tend to be higher. This lends further credence to the arguments of Kopczuk (2005) and how the tax base affects the value of the E.T.I.

Holmlund & Soderstrom (2011) introduced a forth model for making long-term predictions, which has to date not been widely used.

They specify a regression model as

$$y_{it} = \alpha + \beta_1 n_{it} + \beta_2 n_{it-1} + X_{it}\gamma_1 + X_{it-1}\gamma_2 + \rho y_{it-1} + \theta_i + \varepsilon_{it} \qquad \text{(A5)}$$

Where:
y_{it} is log income.
$n_{it} \equiv \ln(1-\tau_{it})$ is the log of the net-of-tax rate.
τ_{it} is the marginal tax rate.
X_{it} is a vector of other regressors.
θ_i is an individual-specific fixed effect.
ε_{it} is a random error term with mean value 0.
i indexes individuals.
t indexes years.

The short-run E.T.I. is given by β_1 and the long-run E.T.I. is given by $(\beta_1 + \beta_2)/(1-\rho)$, where $\rho \in (0, 1)$ is thought to hold. This last expression is valid when $y_{it} = y_{it\text{-}1}$ in the steady state, which will be reached if there are no other factors which exert influences on reported taxable income and therefore prevent income from becoming stationary.

Some of the source articles include additional variations on estimations techniques for the E.T.I. which warrant further mention:

i) Canada

The research in Canada has focused principally on data at the provincial level, since there have been few changes in federal tax rates.

Gagne, Nadeau & Vaillancourt (2001) use provincial tax rates to calculate very high values for E.T.I.s. They find evidence of very high elasticities of even total income, not just taxable income. They argue that tax revenues can be increased by reducing tax rates (i.e. a Laffer response). This is especially interesting because they find estimates in excess of 1 for even the middle classes, and almost 2 for the higher classes.

Milligan & Smart (2013) find that cross-province tax shifting accounts for about two thirds of the total E.T.I. to unilateral provincial tax changes. They find that almost all of the reaction to changes in tax rates is from the top 1% of the population, with general E.T.I.s in the neighbourhood of 0.5 to 1 for this group. Trusts account for a large portion of this income shifting, rather than physically moving. Their 2014 paper estimates the E.T.I. for the top earning 1% of the population more precisely, with a value of 0.664.

It is interesting to note that the data they used and their identification hypotheses are similar to Gagne, Nadeau & Vaillancourt (2001), but the estimates are much smaller. Part of the reason is likely the different regression tools used in the studies, which for Gagne, Nadeau & Vaillancourt (2001) were ‘Seemingly Unrelated Regression’ (S.U.R.E., as introduced in Zellner (1962)), followed by a Prais-Winsten transformation and a second, iterative S.U.R.E. regression.

Milligan & Smart (2014) instead use a time series regression of the form

$$\log(\sigma_{pt}) = \beta_0 + e \log(1-\tau_{pt}) + \beta_1 \log(TotIncome_{pt}) + \delta_p + \lambda_t + \nu_{pt} \qquad \text{(A6)}$$

Where:
$\log \sigma_{pt}$ is the natural logarithm of the top one percent share of income.
e is the elasticity of interest.
$\log(1-\tau_{pt})$ is the logarithm of the net-of-top-tax rate.
$\log(TotIncome_{pt})$ is the natural logarithm of total provincial income.
δ_p is a set of provincial fixed effects.
λ_t is a set of time period fixed effects.

ν_{pt} is the error term.
p indexes provinces, and *t* indexes years.

This is similar to the model presented in Saez (2004). They also weight the data by provincial population, following techniques introduced in Solon, Haider & Wooldridge (2015).

ii) New Zealand

Carey et al (2015) propose alternative instruments for the net-of-tax rate.

Whereas the afore-mentioned standard instrument involves assuming that income just prior to a tax reform would have remained the same without changes to the tax code, Carey et al propose estimating parameters to model changes to income at various points along the income distribution during years without tax reforms. They then use these parameter estimates to model how pre-tax reform income is expected to behave without any tax changes. The tax rate that would have applied to the individual without changes to the tax code can then be used as an instrument.

They propose a second alternative instrument which estimates an expected tax rate. They begin by calculating the mean and variance of log income for individuals who had a specified income in past years. With estimates of such means and variances, they calculate the probability of an individual being in each of the possible post-reform marginal tax brackets based on her pre-reform income. Each probability is multiplied by the marginal tax rate associated with that bracket. They then add all of the products to obtain the expected tax rate of an individual. 1 minus this expected tax rate can be used as an instrument for the marginal net-of-tax rate.

B) Selected Data

Table A1 presents the most relevant data used in the regressions of section 5. In particular, we can see that the amounts of dividends received in the 1990s, even in nominal terms, are somewhat flat compared with other periods. 1997 is an obvious exception, though this was not close to years where the inclusion rate on capital gains changed. This could explain the counter-intuitive results of the relationship between dividends and lags of the capital gains inclusion rate found in section 5.

Table A1

Reference Totals and Percentage Changes for Total Amounts of Investments Received Nationally and Inflation - Canada

Year	Interest (Billions CDN $)	Interest Increase (%)	Dividends (Billions CDN $)	Dividends Increase (%)	Capital Gains (Billions CDN $)	Capital Gains Increase (%)	Inflation (%)
1972	1.1	0.0	0.5	32.4	0.3	15.7	5.0
1973	1.8	63.6	0.6	22.2	0.5	6.3	7.5
1974	2.7	50.0	0.6	7.3	0.6	-14.7	11.0
1975	2.5	-7.4	0.7	10.2	0.8	-2.3	10.7
1976	2.3	-8.0	0.6	-12.3	1.3	2.4	7.5
1977	2.6	13.0	0.7	19.3	1.5	-3.4	8.0
1978	3.6	38.5	1.1	61.8	2.4	14.3	9.0
1979	5.9	63.9	1.2	9.1	4.1	34.4	9.1
1980	6.2	5.1	1.3	8.3	5.5	34.9	10.1
1981	9.1	46.8	2.3	76.9	4.7	3.4	12.5
1982	9.6	5.5	2.4	4.3	2.7	-25.0	10.8
1983	8.0	-16.7	2.5	4.2	4.1	44.4	5.9
1984	9.6	20.0	3.1	24.0	4.0	0.0	4.3
1985	8.7	-9.4	3.9	25.8	5.7	14.4	4.0
1986	8.0	-8.0	4.1	5.1	9.0	11.2	4.2
1987	8.2	2.5	4.0	-2.4	14.6	20.2	4.4
1988	8.4	2.4	7.3	82.5	12.8	-8.7	4.0
1989	9.9	17.9	4.9	-32.9	17.0	15.1	5.0
1990	12.0	21.2	4.8	-2.0	11.1	-9.6	4.8
1991	10.0	-16.7	5.6	16.7	10.8	1.4	5.6
1992	8.3	-17.0	5.4	-3.6	12.0	-1.7	1.5
1993	7.7	-7.2	7.0	29.6	19.0	14.5	1.9
1994	10.0	29.9	6.6	-5.7	18.2	9.9	0.2
1995	14.0	40.0	7.1	7.6	10.0	3.4	2.1
1996	14.0	0.0	7.9	11.3	13.1	18.8	1.6
1997	15.0	7.1	12.0	51.9	17.5	22.7	1.6
1998	15.0	0.0	12.0	0.0	17.9	4.7	1.0
1999	14.0	-6.7	11.0	-8.3	19.0	4.5	1.7

2000	17.0	21.4	12.0	9.1	29.8	37.3	2.7
2001	14.0	-17.6	9.7	-19.2	16.2	-19.8	2.5
2002	9.6	-31.4	13.0	34.0	13.7	-9.2	2.3
2003	8.4	-12.5	12.0	-7.7	16.8	0.9	2.8
2004	9.2	9.5	14.0	16.7	22.5	20.7	1.9
2005	13.0	41.3	19.0	35.7	31.8	17.6	2.2
2006	19.0	46.2	28.0	47.4	38.6	18.4	2.0
2007	21.0	10.5	31.0	10.7	43.9	13.9	2.1
2008	19.0	-9.5	32.0	3.2	24.8	-8.4	2.4
2009	14.0	-26.3	26.0	-18.8	20.8	-18.6	0.3
2010	13.0	-7.1	28.0	7.7	29.3	18.8	1.8
2011	13.0	0.0	29.0	3.6	31.0	7.4	2.9
2012	15.0	15.4	29.0	0.0	28.1	-6.7	1.5
2013	16.0	6.7	32.0	10.3	32.8	5.4	0.9

Note: Data source was in Canadian dollars rounded to the nearest billion, meaning that 0% changes indicate only that there were no differences in the data for the two years in question, not that there was 0 growth in the amounts received.

Tables A2 and A3 present the migration and tax data used in the regressions of section 6. Table A4 presents corresponding G.D.P. data.

Table A2

Migration Rates between Nordic Countries

Year	Sweden Denmark Immigration	Sweden Denmark Emigration	Sweden Norway Immigration	Sweden Norway Emigration
1956	6572	2234	2335	1450
1957	5599	2434	2291	1444
1958	3661	2791	1752	1423
1959	2492	3456	1708	1283
1960	2695	3316	2281	1187
1961	2695	3115	2364	1293
1962	2411	2686	2090	1451
1963	2990	2160	2155	1397
1964	2783	2463	2582	1449
1965	2831	2150	3600	1572
1966	2556	2386	3158	2037
1967	1869	1986	2072	2177
1968	2725	2173	2210	2397
1969	3261	2224	2650	1924
1970	3609	3397	2837	2602
1971	2431	3403	2258	2969
1972	2126	3070	1996	3209
1973	2357	3383	2286	2671
1974	7142	2134	2515	2414
1975	11507	4535	2815	2237
1976	3996	4912	2639	2194
1977	2572	4154	2180	2498
1978	2000	3577	1749	2482
1979	2134	2890	2118	2085
1980	2521	2366	2018	2268
1981	1881	2248	1661	2412
1982	1635	2215	1766	2536
1983	1578	1797	2501	2159
1984	1654	1912	3122	2118
1985	1742	1914	2627	2550
1986	1820	1930	2852	3180
1987	1816	1730	3416	3738
1988	2952	1688	4905	3823
1989	5081	2459	10623	3170
1990	3719	3142	8620	5016

1991	2226	3026	3339	5271
1992	1803	2574	2643	4688
1993	1718	2357	2126	4677
1994	2379	2377	2562	4388
1995	2354	2628	2674	4383
1996	2021	2742	2589	4908
1997	1627	2671	3015	7093
1998	1927	2445	4293	7765
1999	2194	2196	5496	5912
2000	2996	2084	6577	4955
2001	3564	2247	6287	4421
2002	4250	2241	6374	4404
2003	4603	2585	5807	4391
2004	4674	3024	4884	4211
2005	5008	3341	4317	4445
2006	6432	3456	4489	5018
2007	6615	4307	4714	6006
2008	5605	4754	5055	7206
2009	5457	5144	5097	7188
2010	4962	5195	6372	8846
2011	4583	4856	7455	9753
2012	4011	4471	6705	7379
2013	3676	5015	6220	6934

Table A3

Marginal Tax Rates across Nordic Countries

Year	Swed. lower tax rate	Swed. middle tax rate	Swed. top tax rate	Nor. Lower tax rate	Nor. Middle tax rate	Nor. top tax rate	Denmark top tax rate
52				42.1	50.2	69.5	
53				42.4	51.1	70.9	
54				42.7	51.4	73.1	
55							
56	29.1	32.9	41.3				
57	29.3	33.5	40.6	43.0	54.1	76.4	
58	30.1	35.3	41.3	43.0	56.3	78.2	
59	31.6	38.2	41.7	40.8	47.4	76.0	
60	32.0	38.5	41.9	40.8	47.6	76.0	
61	32.3	38.8	45.6	40.8	49.4	76.1	
62	34.9	39.0	45.8	41.3	53.4	76.9	
63	35.1	39.1	50.1	44.6	55.3	77.2	
64	35.9	43.9	50.7	45.3	55.3	77.2	50.4
65	36.4	42.1	51.2	41.3	51.3	73.2	50.6
66	38.8	42.7	52.9	41.2	41.2	73.2	
67	43.3	46.4	53.4	45.2	45.2	73.2	
68	44.1	47.1	54.0	46.3	48.3	69.2	61.0
69	44.7	47.7	55.3	34.8	41.2	64.5	
70	45.2	48.2	55.8	31.0	37.2	53.5	
71	35.9	47.3	60.6	35.0	42.0	54.0	62.0
72	42.8	57.7	61.8	36.0	45.0	54.1	64.0
73	40.1	62.3	61.9	37.0	54.0	56.0	65.0
74	43.9	63.2	62.0	40.0	50.0	68.2	64.0
75	47.2	58.2	73.2	38.0	43.0	69.7	65.0
76	48.2	64.2	75.2	37.0	48.0	69.4	66.0
77	41.9	62.9	75.9	30.0	41.0	69.4	67.0
78	41.7	59.7	77.7	32.0	38.0	73.4	67.0
79	45.0	62.0	78.0	38.0	48.0	73.4	68.0
80	43.1	59.1	82.1	38.0	48.0	72.8	69.0
81	43.6	55.6	82.6	32.0	38.0	70.7	69.3
82	43.7	58.7	82.7	32.3	38.3	70.8	69.0
83	40.2	53.2	75.2	33.3	37.3	69.3	74.3
84	37.3	53.3	70.3	33.3	36.8	63.8	72.5
85	34.4	50.4	65.4	33.2	36.2	61.9	71.8
86	45.3	50.3	70.3	33.8	36.8	68.8	71.5
87	43.4	50.4	70.4	33.4	36.4	67.4	72.1

88	50.6	50.6	75.6	34.2	44.2	63.2	72.6
89	47.8	47.8	72.8	34.5	34.5	62.0	72.9
90	41.2	55.2	66.2	33.8	43.8	59.3	72.6
91	34.3	34.3	51.2	34.3	41.8	57.8	72.6
92	34.1	34.1	51.0	30.2	35.8	49.5	73.2
93	34.8	34.8	51.0	30.2	35.8	49.5	73.2
94	35.5	35.5	51.1	30.2	35.8	49.5	69.7
95	37.4	37.4	56.5	30.2	35.8	49.5	68.1
96	38.2	35.0	56.7	30.2	35.8	49.5	66.0
97	38.9	35.7	56.7	30.2	35.8	49.5	65.8
98	38.6	35.6	55.8	30.2	35.8	49.5	64.2
99	39.7	36.6	50.6	29.9	35.8	49.3	64.3
00	38.3	35.2	50.4	35.8	35.8	49.3	63.3
01	37.2	34.2	50.5	35.8	35.8	55.3	62.9
02	36.0	32.9	50.5	29.4	35.8	55.3	62.4
03	35.5	35.5	51.2	29.1	35.8	55.3	62.3
04	35.9	35.9	51.5	29.1	35.8	55.3	63.0
05	35.4	35.4	51.6	27.1	35.8	51.3	63.0
06	34.8	34.8	51.6	26.3	35.8	47.8	63.0
07	31.6	31.6	51.6	26.3	35.8	47.8	63.0
08	30.4	30.4	51.4	26.3	35.8	47.8	63.0
09	29.5	29.5	51.5	26.3	35.8	47.8	62.8
10	28.6	28.6	51.6	26.3	35.8	47.8	56.1
11	28.6	28.6	51.6	26.3	35.8	47.8	56.1
12	28.6	28.6	51.6	26.3	35.8	47.8	56.1
13	28.7	31.7	51.7	26.3	35.8	47.8	56.2

Note: Blank spaces represent years where data were not available.

Table A4 provides data used on GDP growth rates in the Nordic nations.

Table A4

GDP Growth Rates across Nordic Countries

Year	Denmark GDP Growth	Norway GDP Growth	Sweden GDP Growth
1956		12.8	3.2
1957		6.8	2.8
1958		0.4	2.1
1959		6.4	4.1
1960		-2.6	5.4
1961	6.4	6.3	5.7
1962	5.7	2.8	4.3
1963	0.6	3.8	5.3
1964	9.3	5.0	6.8
1965	4.6	5.3	3.8
1966	2.7	3.8	2.1
1967	3.4	6.3	3.4
1968	4.0	2.3	3.6
1969	6.3	4.5	5.0
1970	2.0	2.0	6.7
1971	2.7	4.6	0.6
1972	5.3	5.2	2.4
1973	3.6	4.1	4.2
1974	-0.9	5.2	3.2
1975	-0.7	4.2	2.7
1976	6.5	6.8	0.4
1977	1.6	3.6	-1.5
1978	1.5	4.9	2.0
1979	3.5	4.4	4.0
1980	-0.4	4.6	1.4
1981	-0.9	1.6	-0.2
1982	3.0	0.2	1.2
1983	2.5	4.0	1.9
1984	4.4	6.1	4.3
1985	4.3	5.6	2.2
1986	3.6	4.0	2.8
1987	0.3	1.8	3.4
1988	1.2	-0.3	2.6
1989	0.2	1.0	2.7
1990	1.0	1.9	1.0
1991	1.3	3.1	-1.1
1992	2.0	3.6	-1.2

1993	-0.1	2.8	-2.0
1994	5.5	5.1	4.1
1995	3.1	4.2	4.0
1996	2.9	5.0	1.5
1997	3.3	5.3	2.9
1998	2.2	2.6	4.2
1999	2.9	2.0	4.5
2000	3.7	3.2	4.7
2001	0.8	2.1	1.6
2002	0.5	1.4	2.1
2003	0.4	0.9	2.4
2004	2.6	4.0	4.3
2005	2.4	2.6	2.8
2006	3.8	2.4	4.7
2007	0.8	2.9	3.4
2008	-0.7	0.4	-0.6
2009	-5.1	-1.6	-5.2
2010	1.6	0.6	6.0
2011	1.2	1.0	2.7
2012	-0.7	2.7	-0.3
2013	-0.5	0.7	1.2
2014	1.1	2.2	2.3

Note: Blank spaces represent years where data were not available.