VATT Publications 68

Essays on behavioral responses to income taxes

Tuomas Matikka

GOVERNMENT INSTITUTE FOR ECONOMIC RESEARCH

VATT PUBLICATIONS

68

Essays on behavioral responses to income taxes

Tuomas Matikka

Valtion taloudellinen tutkimuskeskus
Government Institute for Economic Research
Helsinki 2014

Tuomas Matikka, Government Institute for Economic Research

ISBN 978-952-274-126-4 (nid.)
ISBN 978-952-274-127-1 (PDF)

ISSN 0788-4990 (nid.)
ISSN 1795-3332 (PDF)

Valtion taloudellinen tutkimuskeskus
Government Institute for Economic Research
Arkadiankatu 7, 00100 Helsinki

Edita Prima Oy
Helsinki, December 2014

Essays on behavioral responses to income taxes

Tuomas Matikka

Academic dissertation to be presented, by permission of the Board of the University of Tampere School of Management, for public examination at the Paavo Koli sali (Pinni A), Kanslerinrinne 1, on December 5, 2014, at 12 a.m.

Abstract

This dissertation consists of an introductory chapter and four empirical essays on behavioral responses to income tax incentives. The introductory chapter presents the conceptual background of the welfare analysis of income taxation, highlighting the crucial roles of behavioral elasticities and the structure of the behavioral response. In addition, I present an overview of recent empirical literature on behavioral responses to income taxes.

The first essay studies the extent and significance of income-shifting behavior among the owners of privately held corporations in Finland. The dual income tax system in Finland offers considerable incentives and opportunities for income-shifting between wages and dividends for business owners. Extensive panel data at both the owner and firm level together with Finland's dividend tax reform of 2005 enable us to credibly analyze the empirical significance of income-shifting. The results show that income-shifting is evident. The income-shifting response is homogenous across different owners and firms. However, the size of the income-shifting incentive affects the extent of income-shifting activity, indicating that the costs of income-shifting are non-negligible.

The second essay focuses on distinguishing between real responses and income-shifting. Previous literature shows that income taxes affect the behavior of high-income earners and business owners. However, it is still unclear how much of the response is caused by real economic activity (labor supply, effort etc.), and how much is due to income-shifting and other tax-planning activities. This distinction is important because it can greatly affect the welfare implications. We build a model that formalizes the analysis of the elasticity of taxable income under income-shifting possibilities. We show that income-shifting accounts for over two thirds of the overall behavioral response among the owners of privately held corporations in Finland. This considerable income-shifting response halves the marginal excess burden of income taxes among Finnish

business owners, compared to the traditional framework where the overall behavioral response determines the welfare loss.

The third essay studies the elasticity of taxable income (ETI) in Finland. To identify the average ETI, I use changes in flat municipal income tax rates as an instrument for changes in the overall marginal tax rates. This instrument is not a function of individual income, which is the basis for an exogenous tax rate instrument. In general, instruments based on predicted tax rate changes as used in many previous studies do not have this feature. In addition, I estimate behavioral responses for the subcomponents of taxable income, such as working hours, fringe benefits and tax deductions. This structure of the overall ETI has rarely been studied in previous literature. My preferred estimate of ETI in Finland is 0.27, which indicates moderate but non-negligible welfare effects. The subcomponent analysis suggests that working hours and wage rates respond less than tax deductions and irregular forms of compensation such as fringe benefits. This implies that the overall behavioral response is not driven by changes in deep individual utility arguments, such as the opportunity cost of working.

The fourth essay analyzes the role of different frictions in explaining different behavioral responses to similar tax incentives. Recent literature shows that optimization frictions attenuate behavioral responses to tax incentives, compared to the underlying long-run structural elasticity. However, the literature has not addressed the role of different frictions in explaining heterogeneous responses to tax incentives. To study this issue, we analyze to what extent and in what manner taxpayers respond to different tax incentives. We compare the behavioral effects induced by different tax and transfer incentives and institutions within similar or even the same individuals in Finland. We use a local estimation approach and the bunching method in order to produce clear and comparable evidence. We find that taxpayers do not respond at all to small changes in tax incentives, but do respond to large and salient incentives. Moreover, the patterns

of response indicate that some taxpayers are unable to respond even to large incentives, and some are unaware of the tax incentives.

Keywords: income taxation, behavioral responses, individuals

Tiivistelmä

Tämä väitöskirja sisältää johdantoluvun ja neljä itsenäistä empiiristä artikkelia verotuksen vaikutuksesta yksilöiden käyttäytymiseen. Johdantoluku esittelee verotuksen taloudellista tehokkuutta käsittelevää teoreettista kirjallisuutta. Lisäksi johdantoluvussa esitellään keskeisimmät empiiriset tulokset tuloverotuksen käyttäytymisvaikutuksista.

Väitöskirjan kaksi ensimmäistä artikkelia käsittelevät yrittäjien käyttäytymisvaikutuksia. Ensimmäisessä artikkelissa tarkastellaan yrittäjien tulonmuuntoa. Verokannustimista johtuva tulonmuunto vähentää verotuloja sekä heikentää verotuksen taloudellista tehokkuutta. Tutkimuksessa tarkastellaan vuoden 2005 yritys- ja osinkoverouudistuksen vaikutuksia. Uudistus kannusti listaamattomien osakeyhtiöiden omistajia maksamaan itselleen mieluummin palkkaa kuin nostamaan osinkoja. Tulosten perusteella verokannustimien muutos vaikutti selvästi omistajien käyttäytymiseen, ja tulonmuuntoaktiivisuus oli hyvin samanlaista eri omistajaryhmissä ja erilaisissa yhtiöissä. Lisäksi havaitaan, että verokannustimen koko vaikutti tulonmuuntoon palkka- ja osinkotulojen välillä.

Toisessa artikkelissa erotellaan tuloverotuksen aiheuttamat reaaliset käyttäytymisvaikutukset tulonmuunnon aiheuttamista vaikutuksista. Aikaisempien tutkimusten perusteella verotus vaikuttaa selvästi yrittäjien käyttäytymiseen. On kuitenkin epäselvää, kuinka suuri osuus kokonaisvaikutuksesta aiheutuu reaalitaloudellisista päätöksistä ja kuinka suuri osa verosuunnittelusta. Erilaisten käyttäytymisvaikutusten erottelu on tärkeää, sillä erilaiset tavat reagoida verokannustimiin vaikuttavat arvioihin verotuksen taloudellisesta tehokkuudesta. Artikkelissa tutkitaan listaamattomien osakeyhtiöiden omistajien käyttäytymistä Suomessa. Tulosten perusteella yli kaksi kolmasosaa verotuksen aiheuttamasta käyttäytymisvaikutuksesta on selitettävissä tulonmuunnolla.

Huomattava tulonmuuntovaikutus pienentää merkittävästi arviota verotuksen aiheuttamasta taloudellisesta tehokkuustappiosta.

Kolmannessa artikkelissa tutkitaan verotettavan tulon joustoa Suomessa. Verotettavan tulon jousto mittaa sitä, kuinka monta prosenttia veropohja muuttuu, kun yhdestä lisäeurosta käteen jäävä osuus muuttuu yhden prosentin. Verotettavan tulon jouston avulla voidaan kokonaisvaltaisesti arvioida tuloverotuksen aiheuttamaa tehokkuustappiota. Korkeampi tuloveroaste voi esimerkiksi vähentää tehtyjä työtunteja sekä lisätä verosuunnittelua. Verotettavan tulon jousto huomioi eri kanavat, joilla tuloveroihin voidaan reagoida. Tutkimuksessa hyödynnetään kunnallisverotuksessa tapahtuneita muutoksia. Kunnallisverotus tarjoaa hyvän vertailuasetelman erilaisten henkilöiden välille, sillä kunnallisveroprosentti ei riipu henkilön tuloista. Lisäksi kunnallisveroprosentin muutokset muuttavat veroasteita kaikissa tuloluokissa. Tulosten perusteella verotettavan tulon jousto on Suomessa keskimäärin 0,27. Tuloveroprosentin vaikutus veropohjan kokoon on tilastollisesti merkitsevä, mutta tuloverotuksen aiheuttama taloudellinen tehokkuustappio on kuitenkin maltillinen. Tulokset antavat myös viitteitä siitä, että työtunnit ja tuntipalkka reagoivat heikommin veroasteen muutoksiin kuin epäsäännölliset tulot, kuten luontoisedut ja verovähennysten määrä.

Neljännessä artikkelissa tarkastellaan käyttäytymisvaikutuksiin tai niiden puuttumiseen vaikuttavia tekijöitä. Viimeaikaisen tutkimuskirjallisuuden perusteella käyttäytymisvaikutuksiin liittyvät kitkatekijät, kuten jäykät työmarkkinat sekä epätietoisuus verosäännöistä ja -kannustimista, heikentävät veronmaksajien reagointia verokannustimiin. On kuitenkin edelleen epäselvää, miten erityyppiset kitkatekijät selittävät havaittua käyttäytymistä, ja miten erilaiset tekijät vaikuttavat havaitun käyttäytymisvaikutuksen kokoon ja muotoon. Tutkimuksessa vertaillaan erilaisten verokannustimien vaikutusta samankaltaisten tai samojen henkilöiden käyttäytymiseen. Tutkimuksessa hyödynnetään paikallista ja tulojakauman muotoon perustuvaa estimointimenetelmää

(nk. bunching -menetelmä). Menetelmällä saadaan luotettavaa ja vertailukelpoista evidenssiä verokannustimien käyttäytymisvaikutuksista. Tulosten perusteella kitkatekijät vaikuttavat merkittävästi veronmaksajien käyttäytymiseen. Veronmaksajat eivät reagoi pieniin kannustimiin, mutta jos verokannustimet ovat merkittäviä sekä selkeitä ja helppo ymmärtää, niin käyttäytymisvaikutuksia havaitaan. Lisäksi havaitaan, että työmarkkinoihin liittyvät kitkatekijät heikentävät keskimääräisiä käyttäytymisvaikutuksia vaikka verokannustimet olisivat merkittäviäkin.

Asiasanat: verotus, käyttäytymisvaikutukset, yksityishenkilöt

Acknowledgments

The long process of writing this PhD thesis is finally coming to an end, and it is time to express my gratitude to everyone who helped and supported me along the way. First, I would like to thank my supervisor Professor Jukka Pirttilä. His insightful guidance and prudent advice has greatly improved this thesis and considerably increased my understanding of economics in general. I am grateful to Jarkko Harju and Tuomas Kosonen for being co-authors of articles in this thesis and for many fruitful discussions on economics, not to forget the various activities unrelated to research. It has been a privilege to work with both of you, and I am already looking forward to all future research projects.

I thank the two pre-examiners of this thesis, Professor Hannu Tervo and Associate Professor Håkan Selin, for their insightful and perceptive comments which greatly improved the manuscript.

I have been privileged enough to work in many different institutes and interact with many intelligent and insightful people during my graduate studies. First, I would like to thank Professor Matti Tuomala for his encouragement and advice in the beginning of my graduate studies. In addition, I thank other professors and all my colleagues and friends at the University of Tampere for many great conversations both on and off the topic of economics.

I thank Director Seija Ilmakunnas and the former Director, Dr. Jaakko Kiander for the opportunity to visit the Labour Institute for Economic Research for one year. Also, I would like to thank the personnel at the Institute for a creative research environment. I would especially like to thank Olli Kärkkäinen with whom I shared a room with for all the discussions on various aspects of life.

I am grateful to Professor Joel Slemrod for the opportunity to visit the University of Michigan for one year, and for his insightful guidance, advice and comments on my work. In addition, I thank everyone at the Office of Tax Policy Research in Michigan for a warm welcome and great conversations on economics and academic life.

For the last two and a half years I have been privileged to work at the Government Institute for Economic Research, VATT. I am grateful to Research Director Seppo Kari and Dr. Essi Eerola for the opportunity to work at VATT, and for numerous discussions and very insightful comments. I would also like to thank the many colleagues at VATT for all their support, advice and encouragement.

I am especially grateful for the financial support I have received from the Finnish Doctoral Programme in Economics, Yrjö Jahnsson Foundation, Finnish Cultural Foundation, OP-Pohjola Group Research Foundation, Nordic Tax Research Council and Emil Aaltonen Foundation.

I am enormously grateful to my mom Anne and to my dad Jukka for their encouragement and support during this project. I would also like to thank my grandfather Lasse and my grandmother Mari for their endless encouragement. In addition, I would like to express my gratitude to my sister Elina and her husband Sami for all their support, and especially for their endless patience with accommodating me in their apartment for numerous occasions during the first and second year of my PhD studies. In addition, I thank all my friends outside the economics profession for all the support.

Finally, the greatest thanks go to my wonderful wife Marja who unconditionally and tirelessly supported me during the process. You mean the world to me, and I would have not been able to do this without you being there.

I dedicate this book to my son Otto, who has brought a great deal of joy and laughter into our lives.

Helsinki, November 2014

Tuomas Matikka

Contents

CHAPTER 1

Introduction

Throughout the modern history of economics, the literature has defined the features of a good tax system. Perhaps the most essential characteristics are the following: first, a tax system should be *productive* enough in order to provide the appropriate amount of public goods such as street lighting and defense, and other tax-funded goods and services such as health care and primary education. Second, a tax system should be *efficient* and distort individual choices as little as possible, compared to a hypothetical situation with no taxes. Third, a tax system should be *equitable* and generate the desired redistribution of income. Finally, a good tax system should be simple, easy to understand and *inexpensive* to administer.

All of the above features affect one another. The need for larger tax revenue might increase the complexity of a tax system if additional tax instruments are implemented. Also, increasing tax revenue might require an increase in existing tax rates, which increases distortions and decreases efficiency. Furthermore, reinforcing the redistributional goals might contradict economic efficiency, as increasing the progressivity of the tax system might increase behavioral responses at the top of the income distribution. In summary, designing real-life tax systems requires many trade-offs between different objectives, and a good tax system is a functional combination of its characteristics.

One of the focal trade-offs in economic policy lies between equity and efficiency. Many policies that increase equity and the redistribution of income also increase the distortions caused by taxes. Similarly, more efficient tax systems are usually less equitable. Therefore, in order to address this central issue, we need credible empirical evidence of both the effectiveness of equitable policies and the efficiency of real-life tax systems.

How different individuals respond to tax incentives is at the core of public finance and modern tax policy debate. Basic economic theory shows that taxation is more efficient the less individuals respond to taxes. In other words, behavioral responses to taxation define the scope of the welfare loss induced by taxes. Therefore, along with equity considerations, behavioral responses are the key factor to consider when designing income tax schedules.

This doctoral thesis focuses on behavioral responses to income taxes and the efficiency of income taxation. The thesis consists of four essays which analyze these issues from different perspectives.

The first two essays analyze behavioral responses focusing specifically on the inefficiency caused by tax avoidance. The first of these papers studies the extent and significance of income-shifting between tax bases among owners of privately held businesses. The second paper distinguishes between real responses and income-shifting between tax bases, highlighting the importance of separating these effects when evaluating the effectiveness of income taxation among individuals with income-shifting opportunities.

The last two papers focus on the structure and composition of behavioral responses. The third essay studies the effectiveness of income taxation among Finnish income-earners using the elasticity of taxable income (ETI) framework (Feldstein 1995, 1999) and changes in local municipal tax rates. In addition to overall responses, this essay characterizes the significance of the different components of the overall elasticity. The fourth essay studies the mechanisms behind different responses to tax incentives. This essay emphasizes that the observed behavioral responses may be moderate because of different frictions related to optimization behavior, such as unawareness of the details of the tax code and inability to respond to tax incentives due to labor market rigidities.

1.1. Conceptual background

A lump-sum tax is a tax which is levied based on some unchangeable characteristic of an individual. A simple lump-sum tax would be a head tax of the same size for all individuals, regardless of their income levels. By definition, lump-sum taxes do not induce any behavioral responses, and therefore do not distort economic decision-making. However, lump-sum taxes are rarely applied in actual tax systems, as they perform badly in implementing the equity and redistributional aspects of taxation. Also, there are only a few individual characteristics that are truly unchangeable, and it might be difficult for the government to correctly observe them for each individual. Therefore, most real-life tax systems comprise different types of distortive taxes, such as proportional and progressive income taxes in which the tax burden increases along with income. Nevertheless, non-distortive lump-sum taxes provide an intuitive theoretical benchmark when analyzing the excess burden caused by distortive taxes.

Early theoretical studies on the economic efficiency of taxation show that behavioral responses to taxes define the optimal tax rates for distortive taxes. Ramsey (1927) shows in his famous article that the optimal tax rate on a good depends inversely on the elasticity of demand for that good. Thus in order to minimize the excess burden of taxation, tax rates should be lower the more demand responds to changes in relative prices.

However, generalizations of this Ramsey rule do not take equity considerations into account. For example, according to empirical literature, high-income individuals are more responsive to income taxes and have a larger income tax elasticity (Saez, Slemrod and Giertz 2012). This would lead to lower taxation of high incomes, which in general cannot be considered as an equitable policy. In the case of consumption taxes, the Ramsey rule leads to higher taxation of necessity goods such as food. This is again regressive, since low-income individuals spend a larger share of their income on necessities.

The pioneering article by Mirrlees (1971) considers equity aspects as well as asymmetric information in setting out the optimal income tax rate schedule. In short, the Mirrleesian optimal income tax model is built such that a function of individual utilities is maximized subject to the government budget constraint, taking into account the fact that individuals can respond to taxes. Also, in the model the government cannot observe the different ability types of individuals. This creates the baseline information asymmetry problem of income taxation.

Intuitively, social welfare, measured as the function of individual utilities, increases when the distribution of income is more equal. However, larger income tax rates can have a negative impact on working, which decreases social welfare. This labor supply elasticity with respect to income taxes therefore limits the ability of government to set high tax rates.

There are not many general rules that can be derived from the baseline Mirrlees model. The most famous lessons are that the optimal income tax rate schedule is not linear, and that marginal tax rates are always between 0 and 100% (at the intensive margin of labor supply).

After the Mirrlees (1971) article, there have been many studies that have aimed to produce more applicable policy rules for income taxation. Many studies have used simulation-based methods to set out an optimal income tax schedule. This literature is surveyed in Tuomala (1990). Using a similar modeling framework as Mirrlees, the seminal article by Saez (2001) derives optimal income tax rates using behavioral elasticities and the observed shape of the income distribution. This enables him to define the shape of an optimal income tax rate schedule using numerical simulations. Overall, despite the rather complicated structure of optimal income tax models, the literature underlines that policy recommendations with regard to income tax schedules are tightly related to the behavioral responses caused by income tax rates.

Another relevant question is what is the appropriate statistic for measuring the behavioral response and the effectiveness of income taxation. Since the groundbreaking work of Harberger (1964), many studies have estimated the elasticity of labor supply with respect to the income tax rate. In summary, this literature has found only small elasticities for individuals at the intensive margin of labor supply, which has led to the conclusion that income taxation entails only minor efficiency costs for individuals who are already working. On the other hand, estimated labor supply elasticities at the extensive margin are usually larger, which indicates that the elasticity of the choice to participate in the labor force is significant. (Meghir and Phillips 2010).

A significant contribution by Feldstein (1995, 1999) changed the view of small welfare losses at the intensive margin. Feldstein shows that the compensated elasticity of taxable income (ETI) with respect to the net-of-tax rate (one minus the marginal tax rate) is the sufficient statistic to analyze the overall welfare loss of income taxes. Theoretically, in the taxable income model, individuals optimize such that the marginal cost of both creating and reporting taxable income equals the net-of-tax rate.

The intuition in the Feldstein model is that it does not matter *how* individuals change their behavior due to income tax. Changes in labor supply, deduction behavior, tax avoidance, tax evasion etc. all reflect the inefficiency of income taxation, and all of these channels are reflected in the taxable income (gross income minus tax deductions and tax credits) of an individual. I analyze ETI in Finland in the second, third and fourth essay of the thesis.

The taxable income model of Feldstein implicitly assumes that the marginal *social* cost of reducing labor supply or work effort equals the marginal social cost of tax avoidance and tax evasion. Chetty (2009a) argues that this is not the case in practice. For example, in many cases, the costs of tax avoidance and tax planning include transfers to other agents in the economy (e.g. compensations to tax consultants), which creates a fiscal externality that is not included in the Feldstein model. Also, individuals might overestimate the costs of tax avoidance and tax evasion, which results in optimization errors. These are important issues because ETI is shown to be larger among high-income earners and business owners (see e.g. Saez et al. 2012), who, compared to regular wage earners, have more opportunities to legally avoid and illegally evade taxes.

Chetty (2009a) shows that if the real social cost of evasion or avoidance does not equal the net-of-tax rate, the marginal excess burden can be characterized by the weighted average of taxable income elasticity and gross income elasticity. Intuitively, Chetty (2009a) demonstrates that ETI might overestimate the welfare loss of the income tax if a large proportion of the response comes in the form of tax avoidance. In the extreme case, if all of the response comes from tax avoidance activity that involves no real social costs, the deadweight loss equals zero even if the ETI itself is large. I analyze this issues in the second essay of the thesis.

Also, the baseline ETI model does not account for the fact that taxable income is an endogenous parameter itself. The government can affect the definition of taxable income, for example by changing the approved amount of tax deductions or tax-exempt income. Thus by altering the concept of taxable income, the government can affect ETI (Slemrod and Kopczuk 2002, Kopczuk 2005). Therefore, in addition to overall ETI estimates, it is relevant to know what components of taxable income (e.g. labor supply, effort, tax deductions, tax evasion etc.) are responsive to tax rates and which are not. I study the anatomy of ETI in the second and third essay.

Some recent papers also discuss the implications of optimization frictions for the observed elasticity estimates. For example, Chetty (2012) and Chetty et al. (2011) show that the observed elasticities are small because of various frictions affecting individual behavior, such as job search costs and adjustment costs, and the salience of the tax rules. These attenuate the observed elasticities below the underlying responsiveness, which we would observe in the absence of these frictions. This underlying long-run structural elasticity is the parameter of main interest in welfare analysis.

To define the structural elasticity, we would need to estimate the underlying structural tax responsiveness in the absence of frictions, which tends to be challenging. However, in order to know more about the nature and economic implications of these frictions, we would also need to know what types of frictions actually hinder behavioral responses, and whether different frictions cause different patterns of behavioral response in the long run. I analyze this issue by studying different income tax and income transfer systems in Finland in the fourth essay of the thesis.

Finally, there is a somewhat heated debate among economists on the suitability of structural and reduced-form approaches to estimate sufficient parameters for welfare analysis (see e.g. Angrist and Pischke 2010, and Chetty 2009b). A typical structural model defines a thorough theoretical model of economic behavior, and uses it to estimate or calibrate the relevant behavioral parameters. The reduced-form approach focuses on a careful identification of the causal relationship by using a research design that exploits experimental or quasi-experimental variation without setting up a specific economic model of individual behavior.

Both of these approaches have their pros and cons. Structural models can be used to characterize the welfare implications of non-existent tax systems. However, they rely on many critical economic assumptions which may or may not hold empirically. On the other hand, reduced-form analysis can provide more reliable statistics on behavioral responses within a specific tax system. Nevertheless, parameters derived from a well-defined quasi-experimental setup are not necessarily generalizable to large-scale welfare analysis or to other tax systems.

The ETI framework provides a midway between these two approaches. In the literature, ETI is typically estimated using a quasi-experimental identification strategy.

Under rather general structural conditions, the ETI parameter is all we need for welfare analysis. In other words, it is not necessary to estimate a large set of deep primitives in a full structural model in order to evaluate the marginal excess burden of the income tax. Therefore, the ETI approach provides the sufficient statistics for analyzing the efficiency of income taxation (Feldstein 1999, Chetty 2009b).

1.2. Overview of recent empirical literature

The groundbreaking work of Feldstein (1995,1999) initiated a new wave in empirical public finance and the analysis of welfare losses. Empirically, the key issue is that the taxable income of individuals is widely available. Compared to labor supply elasticity, which requires (survey) information on individual working hours, the elasticity of taxable income can be analyzed using register-based tax record data, which tend to be available in most developed countries.

Feldstein analyzed the US tax reform of 1986 using panel data and a difference-in-differences style approach. He finds that tax rate cuts affecting the high-income individuals resulted in a surge in taxable income, resulting in a large ETI of 1-3 among top income earners. Estimates this range imply that income taxation has large welfare costs, even suggesting that decreasing the income tax rate would have immediate positive welfare effects.

Feldstein's contribution gave rise to several subsequent studies that more carefully consider underlying income trends, mean reversion of income and other important econometric issues. By applying net-of-tax rate instruments which account for only the legislative changes in income taxes, along with a larger set of individual control variables, Auten and Carroll (1999) estimate an ETI of 0.55 when studying the same 1986 tax reform as Feldstein. Similarly, Moffit and Wilhelm (2000) estimate the ETI to be in the range of 0.3-1 when using various tax rate instruments and the 1986 tax reform in the US.

One of the most well-known empirical ETI studies, by Gruber and Saez (2002), uses long panel data and a number of US tax reforms to study ETI. They aim at a careful consideration of underlying income trends and mean reversion. Gruber and Saez (2002) estimate an ETI of 0.17 for low-income earners and 0.57 for high-income earners. In addition, they analyze the responsiveness of gross income. They obtain an earnings elasticity of 0.17 for gross income among high-income earners. This suggests that a considerable part of the response stems from itemized deductions rather than changes in real economic activity. Giertz (2010) follows the methods of Gruber and Saez (2002), and estimates an ETI of 0.30 by using more extensive and more recent panel data. In addition, Giertz (2010) underlines that ETI estimates are sensitive to empirical specification and the chosen time period. This partly indicates the difficulty of separating underlying income trends and mean reversion effects from the estimated ETI.

A recent survey by Saez, Slemrod and Giertz (2012) summarizes the empirical ETI literature, along with an extensive methodological discussion of ETI estimation. In particular, they discuss the pros and cons of cross-sectional and panel data estimation approaches. Saez et al. (2012) conclude that cross-sectional approaches might be more robust than panel data methods, especially when using tax rate variation only at a certain part of the income distribution. This is because it is rather difficult to explicitly control for mean reversion and other non-tax-related changes in the income distribution when using panel data.

However, extensive and even population-wide panel data from outside the US enable rich individual-level controlling, which alleviates these concerns. In turn, combining wide-ranging controlling of individual characteristics and the inclusion of multiple tax reforms across the income distribution enhances the usefulness of individual-level panel data in ETI estimation (Kleven and Schultz 2013).

Most empirical ETI studies to date use US data sets. However, some recent studies also focus on other countries and regions. The particular advantages of many recent papers outside the US are the availability of more comprehensive data and richer tax rate variation.

For the Nordic countries, Blomquist and Selin (2010) study ETI in Sweden. They analyze Swedish tax reforms in the 1980s and estimate an ETI of around 0.20 for men and 1 for women. As a methodological contribution, Blomquist and Selin (2010) derive an instrument based on imputed income which is less correlated with transitory income shocks, compared to the widely-used net-of-tax rate instrument based on taxable income in the base period.

Kleven and Schultz (2013) estimate ETI in Denmark using a number of tax reforms and panel data for all Danish taxpayers. They obtain ETI estimates in the range of 0.05-0.3. Similarly, Chetty et al. (2011) obtain only small ETI estimates for Denmark. For Norway, a recent study by Thoresen and Vattø (2013) reports elasticities below 0.1. In a previous paper, Aarbu and Thoresen (2001) reported zero elasticity for Norway. Unfortunately, there are no previous comprehensive ETI studies available for Finland. One exception is Pirttilä and Uusitalo (2005), who approximate the ETI to be around 0.3 in Finland.

In addition to overall responsiveness to income taxes, some previous studies consider the issue of income-shifting between tax bases or between time periods. Gordon and Slemrod (2000) offer an overview of the income-shifting literature, and show evidence of tax-motivated income-shifting between personal and corporate tax bases among corporate owners in the US. In addition, Devereux, Liu and Loretz (2014) find income-shifting between corporate and personal tax bases in the UK. Also, Sivadasan and Slemrod (2008) find that a decrease in the effective tax rate on wages led to a significant increase in managerial wage compensation for partners of partnership firms in India. Goolsbee (2000) reports that a significant part of the behavioral response to the 1993 tax rate increase in the US was due to re-timing of executive compensation.

In the Nordic countries, Pirttilä and Selin (2011) provide evidence of income-shifting between the capital income and earned income tax bases in Finland after the introduction of the dual income tax system in 1993, after which capital income was taxed separately at a flat tax rate. In addition to overall ETI, Kleven and Schultz (2013) estimate cross-tax elasticities between labor and capital income in Denmark. They find statistically significant cross elasticities, which indicates that tax-motivated income-shifting occurs. le Maire and Schjerning (2013) show evidence of significant intertemporal income-shifting among Danish entrepreneurs. Fjaerli and Lund (2001) find that income tax rates and income-shifting largely affect the choice of income type for entrepreneurs in Norway. Similar type evidence is also available for Sweden (see e.g. Alstadsaeter and Jacob 2012).

To my knowledge, there are no papers available which explicitly study how the (estimated) ETI is comprised. Related to this, Blomquist and Selin (2010) find significant wage rate responses using Swedish data. This indicates that effort responses are a significant factor in the overall ETI. Moffit and Wilhelm (2000) find that the increase in top incomes after the 1986 tax rate cut in the US was not accompanied by an increase in working hours, suggesting that the positive ETI is not caused by an increase in labor supply.

In summary, recent ETI estimates based on panel data and quasi-experimental tax rate variation imply that income taxation does induce statistically significant behavioral responses. However, the estimates are well below the initial estimates of Feldstein (1995, 1999). Accordingly, the evidence based on recent studies suggests that the associated welfare effects of income taxation are moderate at most.

In addition, the structure of the response holds potentially important welfare implications. Existing empirical evidence tends to indicate that a significant part of the overall response among high-income earners is due to income-shifting. This implies that income taxation has only a limited effect on real economic activity in this group. Nevertheless, more evidence on the structure of the behavioral response is still needed in order to make more precise policy conclusions about the effects of income taxes. Finally, the long-run effect of income taxes might be more significant economically, although it is difficult to estimate it using quasi-experimental variation in tax rates.

The most recent wave of behavioral tax literature uses kink points and other discontinuities in the tax code to estimate the local behavioral response. Saez (1999, 2010) initiated this literature by showing that the excess mass of individuals at a marginal tax rate kink is proportional to the local ETI at the kink point. In other words, if taxable income elasticity is significant, we should find individuals bunching around the points in the income distribution where the marginal tax rate exhibits a sharp discontinuous jump.

One important advantage of the bunching approach is that the excess mass can be visibly observed from the income distribution. In addition, the bunching method avoids the intractable econometric issues of tax rate instrumentation and counterfactual income trends related to quasi-experimental estimation utilizing tax rate variation over time. However, the bunching method only delivers local estimates, but income taxes might also have an effect further away from the discontinuity point.

Saez (2010) finds no or only small excess bunching for wage earners in the US, which indicates small or zero local taxable income elasticity. However, he finds that the self-employed bunch actively at some (but not all) kink points, which indicates that tax evasion might largely explain the observed bunching response.

Similarly, Bastani and Selin (2014) find no excess mass around the kink points of the Swedish tax schedule for wage earners. However, they find significant bunching for entrepreneurs. In contrast, Chetty et al. (2011) find significant excess mass for wage earners at some but not all kink points of the Danish income tax rate schedule. The bunching effect is pronounced at larger kink points, indicating that the size of the local discontinuity might matter. Nevertheless, the implied elasticities at the kink points are small (<0.1). In common with earlier studies, Chetty et al. (2011) find distinctive bunching behavior among entrepreneurs.

Kleven and Waseem (2013) use local changes in average tax rates to estimate (structural) earnings elasticities in Pakistan. The jump in the average tax rate creates a notch in the budget set, which creates large incentives to bunch just below the notch point. Kleven and Waseem (2013) find clear bunching at these notches, especially for entrepreneurs, but also for wage earners.

Finally, in order to evaluate long-run structural elasticities, some recent papers examine how behavioral frictions affect the observed elasticities. Chetty et al. (2011) show that bunching is more pronounced at large kink points in Denmark. Similarly, using panel data methods, Kleven and Schultz (2013) show that ETI is larger when the change in the net-of-tax rate is larger. This evidence implies that estimates based on large tax rate changes provide a better benchmark for the long-run structural elasticity if behavioral frictions hinder individuals from responding to smaller tax rate changes. In addition, Kleven and Waseem (2013) use tax notches in Pakistan to uncover structural elasticities. They find that the observed elasticities are largely affected by behavioral frictions. However, these studies do not address the actual type of the (observed) friction. In order to deduce the potential implications of different frictions, we need to have an assessment of the composition of frictions affecting the extent and structure of behavioral responses.

1.3. Summary of the essays

1.3.1. Business Owners and Income-Shifting between Tax Bases: Empirical Evidence from a Finnish Tax Reform. Income-shifting is a common example of a tax avoidance channel. Income-shifting between differently taxed tax bases is generally recognized in the literature, but only a few studies offer credible empirical estimates of the extent of it (Gordon and Slemrod 2000, Sivadasan and Slemrod 2008, Pirttilä and Selin 2011).

This paper provides new and intuitive evidence on the extent and significance of income-shifting behavior. We focus on carefully quantifying the extent and significance of income-shifting among the owners of privately held corporations in Finland. In addition to average income-shifting responses, we analyze the heterogeneity of the response among different types of firms and owners. We also study how the costs of income-shifting and the size of the incentive affect the income-shifting response.

In general, income-shifting possibilities and opportunities to avoid taxes are larger among high-income earners and business owners. Income-shifting incentives are especially pronounced within a so-called dual income tax system (DIT), in which wage income and capital income are taxed separately with different tax rate schedules.

In Finland, the owners of privately held corporations can withdraw income from their firm as a combination of wages and dividends, which are taxed with different tax rules and regulations. Thus these owners have ample possibilities to engage in active income-shifting in order minimize tax payments each year.

We use the dividend tax reform of 2005 in Finland as a source of variation in income-shifting incentives. The reform abolished the single taxation of dividend income, and thus increased the dividend tax rates for many owners of privately held corporations. In general, the reform increased the incentives to replace dividend payments with wages. Importantly, the reform changed dividend tax rates differently: for some owners there were only small changes in tax rates whereas some owners faced large changes in income-shifting incentives.

Variation in incentives together with population-wide data both at the owner and firm level enable us to credibly estimate the extent of income-shifting behavior. The extensive data allow us to precisely define the tax-optimal composition of total gross income (wages + dividends) which minimizes tax payments for each owner before and after the reform. Analyzing how changes in the tax-optimal income composition affect the changes in the realized income composition gives us a unique and novel approach to estimate income-shifting responses.

We find that income-shifting behavior is evident among Finnish business owners. Changes in the tax-optimal income composition have a notable and robust effect on changes in the observed income composition. Furthermore, we find that the effect is homogenous between different firms and owners. This implies that income-shifting behavior cannot be explained or predicted by observed characteristics. However, the size of the income-shifting incentive affects the response, as owners with larger incentives are more active in income-shifting. This suggests that costs related to income-shifting are relevant.

1.3.2. The Elasticity of Taxable Income and Income-Shifting: What is "Real" and What is Not? Previous empirical literature shows that income taxes generate significant behavioral responses among high-income earners and business owners. However, it is not clear how much of this overall response is due to changes in real economic behavior such as labor supply and effort, and how much is caused by tax avoidance.

This paper distinguishes between real economic responses and income-shifting between different tax bases. Based on a theoretical framework, we build an empirical model which formalizes the analysis of the elasticity of taxable income (ETI) under income-shifting possibilities. In addition, we discuss how the explicit inclusion of income-shifting affects the welfare analysis of income taxation. As an empirical example, we analyze the behavioral responses among the owners of privately held corporations in Finland. Within the Finnish dual income tax system, owners can withdraw income from their firm as a combination of wages and dividends, which are taxed with separate tax rate schedules and tax rules.

In the standard model of the marginal excess burden, it does not matter how individuals respond to income taxes, as all behavioral responses (labor supply, effort, tax avoidance, tax evasion etc.) indicate similar welfare effects (Feldstein 1999). However, income-shifting might distort this line of thought for at least two reasons. First, the shifted income is usually also taxed. Thus not all of the overall response is a pure deadweight loss. Second, due to potential and likely fiscal externalities, the real social costs associated with income-shifting might be low, which further decreases the efficiency loss. In the extreme case in which the income-shifting costs are mainly payments to other agents in the economy (such as tax consultants), the marginal excess burden largely stems from real responses (Chetty 2009b).

Our results show that over two thirds of the overall ETI of dividend income is due to tax avoidance through income-shifting. For wage income, the only statistically significant response comes through the income-shifting channel.

Our empirical example highlights that welfare evaluations based on overall ETI estimates might be misleading for individuals with income-shifting possibilities. In the case of Finnish business owners, the marginal excess burden of dividend taxes decreases from 0.9 to 0.4 when we account for the fact that the shifted income is also taxed. If income-shifting induces no real social costs, the excess burden further decreases to 0.3.

Finally, in addition to income-shifting responses, we find that dividend taxes also induce real responses. These effects are present even when analyzing broader firm-level income components that are less subject to tax avoidance, such as the turnover of the firm. In contrast, wage taxes appear not to induce any real effects.

1.3.3. Taxable Income Elasticity and the Anatomy of Behavioral Response: Evidence from Finland. This paper studies the elasticity of taxable income (ETI) using Finnish register-based data. In order to identify average ETI, I use changes in flat municipal income tax rates as instruments for the changes in net-of-tax rates. This instrument is not based on individual income, which provides the basis for an exogenous instrument. In addition to average ETI, I study the structure of the response by analyzing how the subcomponents of ETI, such as labor supply and deduction behavior, react to changes in tax rates. This anatomy of the overall response (Slemrod 1996) has rarely been studied in the literature.

Under general conditions, ETI measures the marginal deadweight loss of income taxation (Feldstein 1999). In addition to labor supply responses, ETI covers changes in, for example, effort and productivity, deduction behavior, tax evasion and tax avoidance. All of these margins are (more or less) important when considering the overall efficiency of a tax system.

Nevertheless, detailed knowledge of how different subcomponents of ETI respond is useful when designing the income tax system and the detailed structure of tax reforms, especially from the point of view of minimizing the excess burden of income taxation. In addition, the analysis of different subcomponents provides information on the actual economic nature of the response. It is rather difficult for policymakers to influence deep individual utility arguments, such as the opportunity cost of working. However, for example, it is easier to influence tax deduction behavior through minor adjustments to tax regulations.

Recent literature highlights that frequently used predicted net-of-tax rate instruments are not necessarily consistent (see Blomquist and Selin 2010 and Weber 2014). These instruments are functions of individual income in the base period, and thus possibly endogenous in a model where changes in taxable income are regressed with changes in the instrumented net-of-tax rate.

In this study I use changes in municipal tax rates as instruments. The flat municipal income tax rate is independent of individual income level, which is the basis for an exogenous instrument in the ETI model. In addition, changes in municipal tax rates have an effect on net-of-tax rates throughout the income distribution, and different municipalities have changed their tax rates differently in different years. These improve the identification of the average elasticity parameter, while avoiding some of the usual difficulties in ETI estimation, namely non-tax-related changes in the shape of the income distribution and the mean reversion of income.

My preferred estimate for the average ETI in Finland is 0.27. The preferred empirical specification includes extensive individual and regional controlling. This point estimate implies a marginal excess burden of around 15%, which is in line with most recent ETI studies. Intuitively, this suggests that income taxation in Finland induces non-negligible but not extensive efficiency losses.

The subcomponent analysis suggests that working hours and wage rates respond less than tax deductions and irregular forms of compensation. This tentatively implies that the overall behavioral response does not stem from profound economic parameters such as the opportunity cost of working. Thus even though the average ETI estimate is not trivial, changing the income tax rate seems to have only a limited effect on labor supply and work effort, especially for full-time workers.

1.3.4. Unwilling, Unable or Unaware? The Role of Different Behavioral Factors in Responding to Tax Incentives. Many empirical studies find varying behavioral responses to similar tax incentives. Recent literature indicates that different optimization frictions could explain the differing responses. Optimization frictions prevent (some) taxpayers from responding to tax incentives in a manner implied by the underlying long-run structural elasticity (Chetty 2012, Kleven and Waseem 2013). The various frictions include, for example, search costs and other labor market rigidities

(Chetty et al. 2011) and knowledge about tax rules and inattention (Chetty, Friedman and Saez 2013, Chetty and Saez 2013). However, previous literature has not systematically addressed what types of institutions cause what kind of frictions, and how different frictions affect the extent and nature of the behavioral response.

In this paper we study to what extent and in what manner individuals respond to different tax incentives. Using detailed and population-wide panel data, we compare behavioral effects induced by different tax and transfer schemes within similar or even the same individuals in Finland. We use the bunching method in order to produce clear, robust and comparable evidence. We study how individuals respond to local changes in marginal income tax rates, and how university students respond to a distinct notch in their budget set caused by an income limit within the national study subsidy program.

We find no bunching at the kink points of the marginal income tax rate schedule. This result holds for any subgroup we study, including the self-employed and university students. In general, the no-bunching result could be due to the small underlying responsiveness, unawareness of the location of the kink points or the concept of marginal tax rates, or an inability to adjust taxable income with reasonable costs.

However, our results suggest that the inability to respond does not fully drive the negligible bunching at kink points. Based on the evidence for self-employed individuals and university students, kink points do not induce responses even though these groups are clearly able to affect their reported taxable income or labor supply in other similar situations (personal *gross* income reported in round numbers by the self-employed, and significant bunching at income notches by students). This suggests that either unawareness or low elasticity could explain the negligible bunching at kink points, at least for the self-employed and students.

Earning income above a sharp income limit results in Finnish university students losing part of their study subsidy. We find that students bunch actively at this distinctive and relatively salient income notch. However, bunching around the notch is not sharp, and many students are located in the dominated region of behavior where they could earn more disposable income by not exceeding the income limit. These observations suggest that the inability to respond affects the overall behavioral response. However, these pieces of evidence imply that individuals react to sufficiently strong tax incentives even when the ability to respond is attenuated.

Overall, our findings support the view that frictions play an important role in explaining observed taxpayer responses. However, in future research it is important to distinguish more specifically between different types of frictions. Different reasons for responding or not responding can entail different welfare implications. For example, if behavioral responses are attenuated by unawareness, it is possible that individuals will not respond in the future either, at least without a notable change in awareness of the regulations. In contrast, if responses are affected by the inability to respond, the underlying long-run structural elasticity might be larger than the observed response if individuals are able to adjust their behavior over time.

Bibliography

[1] Aarbu, K. and T. Thoresen. (2001). Income responses to tax changes – Evidence from the Norwegian tax reform. National Tax Journal, 54(2): 319–334.

[2] Alstadsæter, A. and M. Jacob. (2012). Income shifting in Sweden – An empirical evaluation of the 3:12 rules. Report to the Expert Group on Public Economics (ESO) 2012:4, Swedish Ministry of Finance, Stockholm.

[3] Angrist, J. and J-S. Pischke. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. Journal of Economic Perspectives, 24(2): 3–30.

[4] Auten, G. and C. Carroll. (1999). The effect of income taxes and household behavior. Review of Economics and Statistics, 81(4): 681–693.

[5] Bastani, S. and H. Selin (2014). Bunching and non-bunching at kink points of the Swedish tax schedule. Journal of Public Economics, 109: 36–49.

[6] Blomquist, S. and H. Selin. (2010). Hourly wage rate and taxable labor income responsiveness to changes in marginal tax rates. Journal of Public Economics, 94: 878–889.

[7] Chetty, R. (2012). Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply. Econometrica, 80(3): 969–1018.

[8] Chetty, R. (2009a). Is the taxable income elasticity sufficient to calculate the deadweight loss? The implications of evasion and avoidance. American Economic Journal: Economic Policy, 1(2): 31–52.

[9] Chetty, R. (2009b). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. Annual Review of Economics, 1: 451-488.

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

[11] Chetty, R., J. Friedman and E. Saez. (2013). Using differences in knowledge across neighborhoods to uncover the impacts of the EITC on earnings. American Economic Review, 103(7): 2683–2721.

[12] Chetty, R. and E. Saez. (2013). Teaching the tax code: Earnings responses to an experiment with EITC recipients. American Economic Journal: Applied Economics, 5(1): 1–31.

[13] Devereux, M., L. Liu and S. Loretz. (2014). The elasticity of corporate taxable income: New evidence from UK tax records. American Economic Journal: Economic Policy, 6(2): 19–53.

[14] Feldstein, M. (1999). Tax avoidance and the deadweight loss of the income tax. Review of Economics and Statistics, 81: 674–680.

[15] Feldstein, M. (1995). The effect of marginal tax rates on taxable income: A panel study of the 1986 Tax Reform Act. Journal of Political Economy, 103: 551–572.

[16] Fjaerli, E. and D. Lund. (2001). The choice between owner's wage and dividends under the dual income tax. Finnish Economic Papers, 14: 104–119.

[17] Giertz, S. (2010). The elasticity of taxable income during the 1990s: New estimates and sensitivity analyses. Southern Economic Journal, 77(2): 406–433.

[18] Goolsbee, A. (2000). What happens when you tax the rich? Evidence from executive compensation. Journal of Political Economy, 108: 352–378.

[19] Gordon, R. and J. Slemrod. (2000). Are 'real' responses to taxes simply income shifting between corporate and personal tax bases?, in J. Slemrod (ed.), Does Atlas Shrug? The Economic Consequences of Taxing the Rich. New York: Harvard University Press and Russell Sage Foundation, 2000, 240–280.

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

[21] Harberger, A. (1964). Taxation, resource allocation, and welfare., in J. Due (ed.), The role of direct and indirect taxes in the federal revenue system. Princeton: Princeton University Press, 1964, 25–70.

[22] Kleven, H. and E. Schultz. (2013). Estimating taxable income responses using Danish tax reforms. Forthcoming American Economic Journal: Economic Policy.

[23] Kleven, H. and M. Waseem (2013). Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan. Quarterly Journal of Economics, 128: 699–723.

[24] Kopczuk, W. (2005). Tax bases, tax rates and the elasticity of reported income. Journal of Public Economics, 89: 2093–2119.

[25] le Maire, D. and B. Schjerning. (2013). Tax bunching, income shifting and self-employment. Journal of Public Economics, 107: 1–18.

[26] Meghir, C. and D. Phillips. (2010). Labour supply and taxes. in S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gummie, P. Johnson, G. Myles and J. Poterba (eds.). Dimensions of tax design: The Mirrlees review. Oxford, Oxford University Press, 2010, 202–274.

[27] Mirrlees, J. (1971). An exploration in the theory of optimum income taxation. Review of Economic Studies, 38(114): 175–208.

[28] Moffit, R. and M. Wilhelm. (2000). Taxation and the labor supply decisions of the affluent. In J. Slemrod (ed.), Does Atlas Shrug? The Economic Consequences of Taxing the Rich. New York: Harvard University Press and Russell Sage Foundation, 2000, 193–234.

[29] Pirttilä, J. and H. Selin. (2011). Income shifting within a dual income tax system: Evidence from the Finnish tax reform of 1993. Scandinavian Journal of Economics, 113(1): 120–144.

[30] Pirttilä, J. and R. Uusitalo. (2005). Veronalennukset vähentävät verotuloja. Kansantaloudellinen aikakauskirja, 4/2005 (in Finnish).

[31] Ramsey, F.P. (1927). A contribution to the theory of taxation. Economic Journal, 37: 47–61.

[32] Saez, E. (2010). Do taxpayers bunch at kink points? American Economic Journal: Economic Policy, 2(3): 180–212.

[33] Saez, E. (2001). Using elasticities to derive optimal income tax rates. Review of Economic Studies, 68: 205–229.

[34] Saez, E. (1999). Do taxpayers bunch at kink points? NBER Working Paper 7366.

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

[36] Sivadasan, J. and J. Slemrod. (2008). Tax law changes, income-shifting and measured wage inequality: Evidence from India. Journal of Public Economics, 92: 2199–2224.

[37] Slemrod, J. (1996). High income families and the tax changes of the 1980s: The anatomy of behavioral response. In Feldstein M. and J. Poterba (eds). Empirical Foundations of Household Taxation. Chicago: University of Chicago Press, 1996, 169–192.

[38] Slemrod, J. and W. Kopczuk. (2002). The optimal elasticity of taxable income. Journal of Public Economics, 84: 91–112.

[39] Thoresen, T. and T. Vattø. (2013). Validation of structural labor supply model by the elasticity of taxable income. Statistics Norway Research Department Discussion Papers No. 738.

[40] Tuomala, M. (1990). Optimal income tax and redistribution. Oxford: Clarendon Press, 1990.

[41] Weber, C. (2014). Toward obtaining a consistent estimate of the elasticity of taxable income using difference-in-differences. Forthcoming, Journal of Public Economics.

CHAPTER 2

Business Owners and Income-Shifting between Tax Bases: Empirical Evidence from a Finnish Tax Reform[1]

Abstract. This study examines the extent of income-shifting between tax bases among the owners of privately held businesses. The dual income tax system in Finland offers noticeable incentives for income-shifting between wages and dividends for business owners. The dividend tax reform of 2005 enables us to study how this particular form of tax avoidance reacts to an exogenous change in tax rates. Our results support highly active income-shifting behavior. We find that the income-shifting effect is homogeneous across different owners and firms. However, we find that the size of the tax incentive affects the size of the response, indicating that costs related to income-shifting are important.

Keywords: business owners, income-shifting, income taxation
JEL codes: H24; H25; H32

2.1. Introduction

In many countries business owners and entrepreneurs have a range of opportunities to shift income between different tax bases in order to reduce overall tax payments. Therefore, it is important to know how income-shifting incentives affect the extent of this behavior. Income-shifting is generally recognized, but only a few studies offer empirical estimates of its size (Gordon and Slemrod (2000); Devereux et al. (2014); Sivadasan and Slemrod (2008)). Our aim is to provide clear and intuitive evidence on the extent and significance of income-shifting behavior.

[1]This essay is joint work with Jarkko Harju (Government Institute for Economic Research VATT). An earlier version of this paper is one of the chapters of the academic dissertation of Jarkko Harju, published in the Government Institute for Economic Research Publications series, 64, November 2013. Another version of this paper is published in the Government Institute for Economic Research Working Papers series, 43, December 2012.

In general, income-shifting is especially relevant for the owners of privately held businesses. Compared to wage earners, business owners have a wider scope of legal possibilities to engage in income-shifting, as they can more easily apply different types of income as a source of personal compensation.[2] Income-shifting possibilities and tax incentives are pronounced within a so-called dual income tax system (DIT). In a typical DIT, the marginal tax rate schedules for labor income and capital income differ significantly from one another.

In this study we carefully quantify the extent and significance of income-shifting between different tax bases among the owners of privately held corporations in Finland.[3] In addition, we analyze the heterogeneity of income-shifting among different types of firms and owners. We also study how the size of the incentive affects the income-shifting response.

We exploit the extensive corporate and dividend tax reform of 2005 in Finland as a source of tax rate variation. The reform increased marginal tax rates on dividends by abolishing the single taxation of dividends. Thus the reform increased incentives to pay wages instead of dividends as a form of personal compensation for many business owners. Importantly, income-shifting incentives changed differently among the owners. For some owners there were only small changes in tax rates, whereas some owners faced large changes in income-shifting incentives.

Total tax record data from the Finnish Tax Administration and the opportunity to link tax record information from the owner-level to the firm-level create an interesting starting point to analyze income-shifting responses. The extensive data allow us to

[2]In addition to many tax bases, income-shifting can also occur in other forms. A well-known example is intertemporal income-shifting, for example, in the form of anticipating the forthcoming tax rate change (see for example Goolsbee (2000) and le Maire and Schjerning (2013)). This paper focuses on the longer run effect of income-shifting between tax bases.

[3]Privately held corporations are defined as corporations that are not listed on a public stock exchange. In the Finnish tax system, dividends from listed and privately owned corporations are taxed at different tax rates and tax regulations.

precisely define the tax-optimal composition of total gross income for each owner before and after the reform. Analyzing how changes in the tax-optimal income composition affect the changes in the realized income composition gives us a unique and novel approach to estimate income-shifting responses. Combining this approach with the variation stemming from the tax reform of 2005 offers us credible empirical evidence on the extent of income-shifting behavior.

We find clear support for the view that business owners are active in income-shifting. Tax-optimal income composition has a clear and robust effect on the realized income composition. However, although changes in income-shifting incentives are large, we do not find a one-to-one income-shifting response with respect to tax incentives. This implies that not all of the owners behave according to a simple tax-minimizing model.

We observe only little heterogeneity in the income-shifting response between different owners or firms. This indicates that income-shifting behavior is not focused on certain types of firms and owners. However, we find that the size of the change in the tax incentive affects the income-shifting responses. This suggests that the costs of income-shifting are important.

Earlier empirical studies concerning income-shifting among business owners have been rather rare. Gordon and Slemrod (2000) offer an overview of the income-shifting literature and show evidence of tax-motivated income-shifting between personal and corporate tax bases among corporate owners in the US. Devereux et al. (2014) also find evidence of active income-shifting between corporate and personal tax bases in the UK. In addition, Sivadasan and Slemrod (2008) find that a decrease in the effective tax rate on wages led to a significant increase in managerial wage compensation for partners of partnership firms in India. In addition, Pirttilä and Selin (2011) show that the relative share of capital income increased among entrepreneurs after the implementation of the Finnish DIT system in 1993. Moreover, concentrated ownership structure is shown to increase tax planning among business owners in the US (Chetty and Saez (2010)).

Within other Nordic Countries, Alstadsæter and Jacob (2012) discuss different tax avoidance channels within the Swedish DIT system, and find evidence of income-shifting between tax bases. Fjaerli and Lund (2001) find support for the hypothesis of active income-shifting among entrepreneurs in Norway. In Denmark, le Maire and Schjerning (2013) provide evidence of income smoothing and intertemporal income-shifting among the self-employed.

The rest of the paper is organized as follows: Section 2.2 presents the institutional background of the Finnish DIT schedule and describes the main attributes of the 2005 tax reform. Section 2.3 depicts the theoretical background for our empirical analysis. Section 2.4 presents the empirical model and descriptive statistics. Section 2.5 presents the results, and Section 2.6 discusses the main findings.

2.2. Finnish dual income tax system and the tax reform of 2005

Since 1993 Finland has applied the principle of Nordic-type dual income taxation (DIT). In DIT, earned income (wages, pensions, fringe benefits etc.) is taxed at a progressive tax rate schedule, whereas personal capital income (interest income, capital gains, dividends from listed corporations etc.) is taxed at a flat tax rate. A distinctive feature of the DIT system is that the flat tax rate on capital income is set much lower than the highest marginal tax rates on earned income. The lower flat tax rate for capital income was motivated for various reasons, for example, broadening the tax base, decreasing the scope for tax arbitrage, and increased global capital mobility which all argue in favor of taxing capital income more leniently.[4]

Within the DIT system, the wide gap between the marginal tax rates on capital income and earned income creates a tricky task for the legislator: How to formalize the taxation of business owners in such a manner that it prevents income-shifting from

[4] A more detailed discussion on the Nordic type DIT can be found for example in Nielsen and Sørensen (1997) and Sørensen (2005).

heavily taxed earned income to more leniently taxed personal capital income? At the same time, the lawmaker needs to assure that the return on invested capital is not overtaxed.

In the Finnish system, this issue is arranged by limiting the amount of flat-taxed dividends. Dividends are split into two parts according to the net assets (assets-liabilities) of the firm. The amount of dividends taxed at the capital income tax rate is based on computational normal rate of return on net assets of the firm. This imputed rate of return (9%) is set to be the same for all owners of privately held corporations. Dividends *less* than the computational normal return are flat-taxed, and any dividends *exceeding* this amount are taxed with the progressive tax rate schedule.[5]

2.2.1. ***The Finnish dual income tax system until 2005.*** Until 2005, Finnish DIT applied a full imputation system of corporate taxes to remove the double taxation of dividends, in which dividend income is taxed both as corporate profits and personal income. In the full imputation system, dividends were exempt from corporate taxes. Thus all dividends were effectively single taxed before 2005. To summarize, taxation of wages and dividends from privately held corporations was organized according to the following rules and principles:

- Dividends:
  - Dividends up to the imputed normal return on the net assets of the firm (assets—liabilities) were subject to the flat personal capital income tax rate of 29%.

[5]For example, with net assets of 400,000 €, the maximum amount of dividends taxed at the flat tax rate is 36,000 € when the imputed return is set to 9% (0.09*400,000=36,000). In other words, any dividends received from the firm below 36,000 € are effectively taxed at the flat tax rate, and any dividends above this amount are subject to progressive taxation with top marginal tax rates above the flat rate. The value of net assets is calculated based the asset and debt values of the firm in the previous year. The individual net asset share of the owner is calculated based on the ownership share of the firm. Also, there are some individual adjustments to the net assets. For example, if the owner or her family members live in a dwelling which is owned by the firm, the value of this dwelling is not included in net assets when calculating the imputed return.

  - Dividends exceeding the imputed normal rate of return were taxed with the progressive tax rate schedule.
  - Corporate taxes were fully credited against the dividend tax liability of a shareholder, resulting in single taxation of both flat taxed and progressively taxed dividends.
- Wages were subject to the progressive tax rate schedule (0-56% in 2002). Wages were single-taxed as they were deductible from firm profits.
- Wages and progressively taxed dividends were not taxed with similar tax rules. Some tax deductions and tax credits were only allowed on wage income. In contrast, progressively taxed dividends were not subject to firm-level social security contributions.[6]

2.2.2. ***The dividend tax reform of 2005.*** From 2005 onward, the full imputation system was abolished, and Finland switched to a system with double taxation of dividends. After the reform, dividends are taxed according to the following principles:

- All dividends became subject to a corporate tax of 26%.
- The splitting rule of dividends according to the imputed rate of return on firm net assets remained unchanged.[7]
- The flat-tax dividends below the imputed return and under 90,000 € remained single-taxed, and are only subject to the flat corporate tax rate of 26%.
- 70% of all other dividend income is taxable in individual taxation, which results in partial double taxation of dividends.
- Wages and progressively taxed dividends are still taxed differently in terms of social security contributions.

[6]Firm-level social security contribution rate is 2-6% of wages, depending on the level of total wages paid and the depreciations made by the firm.
[7]However, the imputed rate of return decreased slightly from 9.6% to 9%.

The taxation of dividend income below the amount corresponding to the imputed return on net assets (9%) did not change significantly in the reform. Effectively, the flat dividend tax rate for dividends below the imputed return and under 90,000 € decreased from 29% to 26%. In contrast, the double taxation rule increased the dividend tax rate for dividends above the imputed return. In addition to individual-level progressive taxation, progressively taxed dividends became subject to the flat corporate tax rate of 26%. Thus after the reform, the minimum effective tax rate for progressively taxed dividends is 26%, compared to 0% before the reform. Furthermore, the flat tax rate increased from 29% to 40.5% for flat-tax dividends over 90,000 €. However, this concerns only a relatively small number of owners.

One important aspect of the reform was its primary motive. According to the European Union Court of Justice, the pre-reform Finnish system of full corporate tax imputation was not in accordance with European Union legislation. Full imputation was granted only to domestic shareholders. Also, the imputed tax credit was not granted to Finnish shareholders whose firms operate abroad. These violated EU regulations on equal tax treatment of all EU citizens. Therefore, Finnish legislators were more or less forced to change the tax system towards a more unified tax treatment. This procedure has important implications for our study. As the reform was not driven by the economic and fiscal conditions in Finland, the tax reform of 2005 can be considered exogenous from the point of view of the owners of privately held corporations.

Finally, the content of the 2005 tax reform was made public already in late 2003. This enabled the owners to anticipate the changes induced by the reform.[8] Also, special transition rules were applied in 2005 to temporarily alleviate the double taxation of dividends. For these reasons, we focus on analyzing the income-shifting effect by using a longer time period of 2002-2008.

[8]For evidence of anticipation effects, see Kari et al. (2008).

2.2.3. ***Tax incentives for income-shifting.*** There are many possibilities for tax avoidance within the Finnish DIT system. For example, the owners of privately held corporations may seek to minimize taxes by dynamically optimizing the level of net assets, and in a static year-to-year context, by choosing an optimal combination of wages and dividends as their personal compensation from the firm. In this paper we focus on the latter case. Importantly, there are only a few minor legal limitations on whether income is withdrawn as wages or dividends from a privately held corporation in Finland.[9]

The tax-optimal division of total income between wages and dividends is relatively complex. The dividend tax rate schedule comprises of both flat-tax and progressive regions, which depend on the net assets of the firm. The amount of flat-tax dividends can be simply calculated based on the net assets position of the firm. However, wage taxes depend on the level of progressively taxed dividends, and vice versa. Wages and progressively taxed dividends are part of the same tax base even though they are effectively taxed with different tax rates. This complicates the optimization process. When optimizing the income composition, the owner needs to simultaneously consider both the effect of net assets and wage income on the tax rate of dividends. We discuss this issue in the light of our empirical analysis in Section 2.4.2.

The dividend tax reform of 2005 changed the income-shifting incentives differently among the owners of privately held corporations. Figure 1 illustrates the changes in income-shifting incentives due to the tax reform of 2005. The Figure presents the marginal tax rates (MTR) on wages and dividends before (2002) and after (2007) the

[9]A corporation cannot distribute dividends more than it holds distributable assets. These include, for example, accumulated profits and non-tied equity. With some firms this might limit the scope for income-shifting. Wages cannot be paid when there is no work contribution to the firm. Otherwise wages may be regarded as a veiled distribution of profits. However, this is a minor issue in our analysis since our sample of corporate owners hold an executive position in the firm, and are thus by default assumed by the tax authorities to work for the firm. In contrast to wages and dividends, other alternatives to withdraw income from the firm are restricted. These include, for example, shareholder loans and share repurchases.

reform with both zero firm-level net assets and with net assets of 170,000 € (median net assets in the data set).[10]

From Figure 1 we can see that wages and dividends were almost equally taxed before the reform for owners with no firm net assets (upper left graph). Differences in tax rates come from the differences in social security payments and tax deductions between wage and dividend income. Dividend taxes increased significantly for this group after 2005 (upper right graph). The double taxation of dividend income increased dividend taxes, making the MTR on dividends higher than the MTR on wages. Thus for the owners with low net assets, the reform induced incentives to shift income from dividends to wages. However, as only 70% of dividends are taxable in individual taxation after the reform, the difference between marginal tax rates decrease at large income levels.

There were no significant changes in the taxation of flat-tax dividends below 90,000€. Before the reform, dividends were in general taxed more leniently than wages for owners with median-level net assets (lower left graph). The reform of 2005 increased dividend taxes for dividends above the flat-taxed region, which brings the MTR on wages and dividends closer to each other (lower right graph).

In addition, the reform did not induce significant changes in income-shifting incentives for owners with very large net assets. However, high-income owners with flat-tax dividends above 90,000 € faced a large change in the MTR on dividends (from 29% to 40.5%). Table A1 in the Appendix presents the marginal tax rates on wages and dividends in numbers for the years 2002 and 2007 and for firm net assets of 0 €, 170,000 € and 1,000,000 €.

[10]Wage tax rates and progressive dividend tax rates include central government taxes, average municipal taxes, applicable individual social security contributions and all automatic deductions and tax credits on either dividend income or wage income or both. In addition, MTR on wages includes firm-level social security contributions. MTR on dividends includes the corporate taxes paid on dividends after the reform.

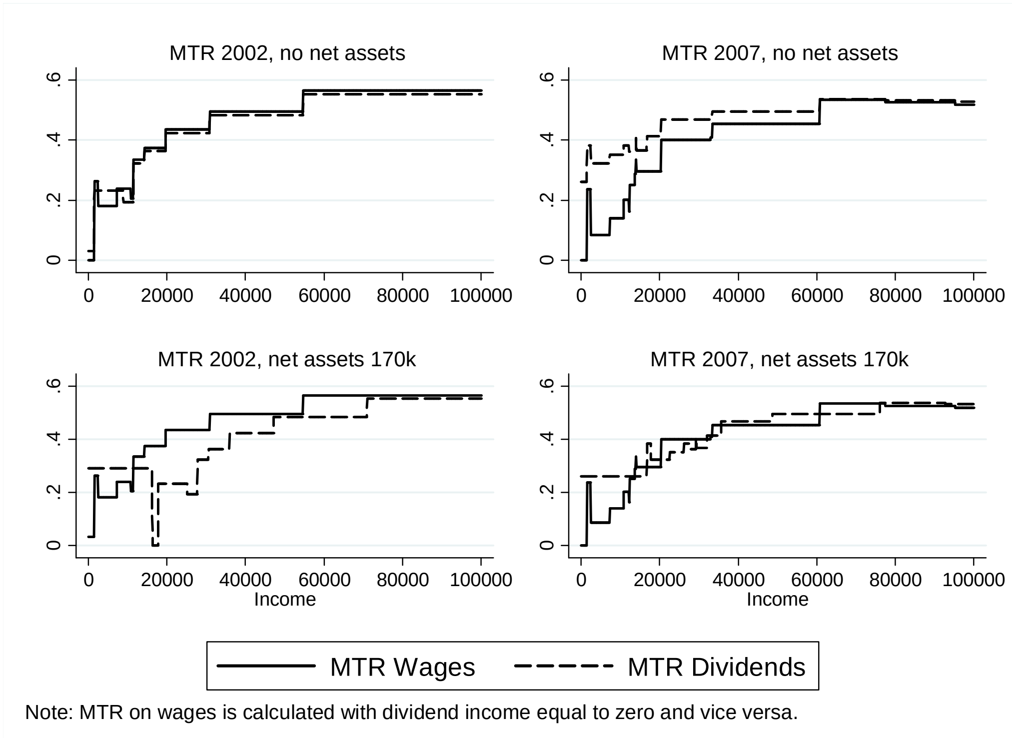


Figure 1. Marginal tax rates (MTR) on wages and dividends: Years 2002 (left) and 2007 (right). Above no net assets, below net assets of 170,000 € (in nominal euros each year)

We do not include mandatory pension and health insurance contributions as a tax on wages in this study. Our empirical analysis is limited to owners who own at least 50% of the firm alone or together with immediate family members, and hold an executive position in the firm. These owners are termed YEL owners in the Finnish tax legislation. YEL owners are subject to special pension insurance rules. YEL owners report a so-called YEL income to the insurance company from which mandatory insurance payments are accumulated from.

Importantly, YEL income does not need to coincide with actual wages paid for the owner. In other words, YEL income can be above or below actual wages paid without implications or sanctions. Thus mandatory insurance contributions have no direct effect on the decision to divide total income into wages and dividends, and are therefore excluded from the income-shifting analysis.[11]

However, annual wages might be correlated with the reported YEL income. Some owners might report the actual wage income withdrawn from the firm as the YEL income. For these owners, pension and health insurance contributions increase or decrease one-to-one with changes in actual wage income. If insurance contributions are regarded as taxes, this reduces the incentives to pay out more wages. Therefore, insurance contributions might mitigate incentives to pay more wages as a response to increased dividend taxation, which would decrease our income-shifting estimate.

2.3. Theoretical framework

The following model depicts the background for our empirical analysis of tax-motivated income-shifting. In the model, the owner of a privately held corporation both owns a significant part of the corporation and works for the firm. We assume that the owner makes all the relevant decisions about the distribution of profits. Distributed profits are paid out to the owner as a combination of wages and dividends. Wages and dividends are taxed at different tax rate schedules.

The owner receives positive utility from her net-of-tax income (i.e. net wages and net dividends) while costs related to income-shifting reduce utility. The utility function is of the form $U(W+D,\gamma)$, where W is net wages and D is net dividends, and γ is the income shifted from wages to dividends. The payout budget constraint is $\Pi - R = W^g + D^g$,

[11]There are regulations for both the lower and upper limits of YEL income, which are, however, also independent of actual taxable wage income. Insurance payments determine pensions when retired, as well as the amount of many income-bound social benefits before retirement (e.g. public health insurance). Thus owners have incentives to report a realistic YEL income which reflects the actual income earning potential. There were no relevant changes in contribution rates or other regulation on insurance payments for YEL owners in the time period we study. The overall average rate of insurance payments on YEL income was 21.1% in 2002 and 20.8% in 2007.

where Π is total profits from the firm before taxes, R is retained earnings, and W^g and D^g are gross wage income and gross dividend income from the firm.

As in Fjaerli and Lund (2001), we focus on the choice of the optimal combination of wages and dividends conditional on given total profits Π and retained earnings R. In other words, we do not model the income-generating process of the firm nor the optimal level of retained and/or distributed profits, and thus simply assume Π and R to be exogenous.[12] We follow this assumption throughout the paper.

More formally, the owner's optimization problem is to maximize

$$U(W+D,\gamma) = (1-t_W)(\tilde{W}^g-\gamma) + (1-t_D)(\tilde{D}^g+\gamma) - \phi(\gamma) \tag{2.3.1}$$

subject to

$$\Pi - R = W^g + D^g \tag{2.3.2}$$

In the model, $\tilde{W}^g$ and $\tilde{D}^g$ represent wage income and dividend income in the absence of income-shifting opportunities. Thus $(\tilde{W}^g-\gamma) = W^g$ is the observed gross wage income, and $(\tilde{D}^g+\gamma) = D^g$ is the observed gross dividend income. $\phi(\gamma)$ denotes the cost of income-shifting, i.e. the cost of changing the tax base. For simplicity, we assume that $\phi(\gamma)$ is convex, smooth, and $\phi(0)=0$ and $\phi'(0)=0$.

[12]The choice of retained earnings (R) is relevant in dynamic tax optimization. R increase net assets, which are the base for determining the flat-taxed dividends in the Finnish DIT system. Other than purely tax-motivated issues also define the amount of R (for example, essential investments and imperfect capital markets). In the analysis, we assume that R is already optimized, or simply taken as given. However, the endogenous nature of R does not change the relevance of the static year-to-year tax minimization problem of choosing the tax-optimal combination of wages and dividends. Also, without year-to-year tax optimization, the benefits from dynamic tax avoidance diminish or vanish altogether.

In equation (2.3.1), $t_W = t_W(W^g, D^g, I)$ and $t_D = t_D(W^g, D^g, I)$ are the average tax rates on wages and dividends, respectively. Both tax rates are always between zero and one. The tax rate on wage income consists of personal income taxes plus firm-level social security contributions. The tax rate on dividends includes dividend taxes plus corporate taxes associated with withdrawn dividends. Wages are assumed to be deductible from firm profits whereas dividends are not. Also, both tax rates depend on income earned outside the firm, denoted by I. This income includes, for example, wages from a secondary job and dividends from other non-listed firms. I is assumed to be exogenous in the model.

Within this general framework, t_W is also a function of dividends, and t_D is a function of wages. This implies that the amount of wages withdrawn from the firm is allowed to have an effect on the tax rate on dividends, and vice versa. Also, we assume that the tax rate schedules of wages and dividends are "well-behaved", smooth and monotonically increasing functions of W^g, D^g and I.

We focus on income-shifting responses with given total income. Thus, to simplify the model, we assume that $\tilde{W}^g$ and $\tilde{D}^g$ are constant. Therefore, we get the optimal income-shifting behavior by taking the first-order condition with respect to γ, which gives us the following result

$$\left(t_W + \frac{\partial t_W}{\partial \gamma}\right) - \left(t_D + \frac{\partial t_D}{\partial \gamma}\right) = \phi'(\gamma) \tag{2.3.3}$$

Equation (2.3.3) says that the combination of gross wages and gross dividends is optimal when the difference between the marginal tax rate on wages $t_W + \frac{\partial t_W}{\partial \gamma} = MTR_W$ and the marginal tax rate on dividends $t_D + \frac{\partial t_D}{\partial \gamma} = MTR_D$ equals the marginal cost of income-shifting.

In our empirical analysis, we relate the change in the observed income combination of wages and dividends (W^g, D^g) to the change in the tax-optimal income combination that minimizes tax payments. Our empirical benchmark of tax-minimizing income-shifting behavior refers to the conceptual case where $\phi'(\gamma) = 0$. With zero income-shifting costs at the margin, the optimality condition is simply

$$MTR_W = MTR_D \tag{2.3.4}$$

Equation (2.3.4) shows that in order to minimize tax payments, owners adjust γ such that the marginal tax rates are equal. This optimality condition determines the combination of gross wages and gross dividends which minimizes taxes. We denote this tax-optimal gross income combination by (W^*, D^*).

In addition, our framework provides an intuitive approach to study the significance of income-shifting costs. These costs might be incurred by, for example, the opportunity cost of time or as payments to tax consultants. If owners choose their income compositions such that $MTR_W = MTR_D$ and thus $(W^g,\ D^g) = (W^*, D^*)$, it would indicate both high tax responsiveness of income-shifting as well as low (marginal) costs of income-shifting. Instead, deviations from the tax-optimal condition imply that marginal income-shifting costs are non-negligible, and that these costs affect the decision on the composition of income. Intuitively, this would suggest that income-shifting is more active when tax incentives are larger.

Finally, assumptions behind the theoretical optimality conditions do not generally hold in practice. First, real-life marginal tax rate schedules are not smooth and continuous. If anything, the schedules are more or less discontinuous piecewise linear functions of income, which implies, for example, that the actual tax-minimizing income combination rarely satisfies the condition $MTR_W = MTR_D$. Second, optimization errors might occur for at least some owners.[13] Nevertheless, comparing actual income combinations to the tax-optimal combination (W^*, D^*) that minimizes tax payments provides the conceptual background for analyzing the extent and significance of purely tax-motivated income-shifting behavior.

2.4. Empirical analysis

2.4.1. *Data.* Our data set comes from the Finnish Tax Administration and it includes information on the financial statements and tax records of Finnish businesses and business owners for the years 2002, 2003, 2007 and 2008.[14] We use it both in a cross-sectional and balanced panel form. The unique characteristic of the data is that they contain basically all Finnish businesses (all public and privately held corporations, partnerships, sole proprietors etc.).

In this study we focus exclusively on the owners of privately held corporations. The data contain all important tax information for the income-shifting analysis, for example, taxable wages and dividends paid to the owner by the firm, and income earned from other sources by the owner. By linking the firm-level and the owner-level data together we can analyze the effects of tax changes on owners' income-shifting behavior while consistently controlling for various firm and individual-level effects.

The owner-level data include only those individuals who received positive dividends from the firm during a tax year. Furthermore, we concentrate only on those owners who work in their own firm in an executive position and own at least 50% of the firm

[13]Also, search costs and other optimization frictions might also matter in optimization behavior (Chetty (2012)). Fjaerli and Lund (2001) suggest that benefits received from paying social security contributions increase wages as a form of compensation, although no compelling evidence has been found to support this view. Also, wages can be seen as a socially more acceptable form of personal compensation. All of these issues imply a deviation from the optimality conditions (2.3.3) and (2.3.4).

[14]As mentioned before, the content of the 2005 tax reform was made public already in late 2003. Kari et. al (2008) show evidence that privately held corporations anticipated the reform by increasing dividend payments right before the reform, and decreasing them right afterward. Therefore, we do not use the years closest to the reform in our baseline analysis in order to alleviate the effects caused by anticipation on the longer-run income-shifting response between tax bases.

alone or together with immediate family members. We discuss the implications of data and sample restrictions in the end of Section 2.5.

2.4.2. *Empirical model.* This section describes the empirical model we use in our analysis. Our aim is to study how the tax-optimal income composition affects the decision to withdraw different types of income from the firm. This relationship can be described with the following cross sectional equation

$$W^g_{i,t} = \beta * W^*_{i,t} + X_{i,t} + C_i + \alpha_t + \varepsilon_{i,t}, \tag{2.4.1}$$

where $W^g_{i,t}$ is realized gross wages from the firm for each owner i in year t. $X_{i,t}$ is a matrix of firm and owner-level variables that affect the amount of gross wage income and the income composition. C_i describes time-invariant variables that affect gross wages, such as the innate ability of the owner.[15] α_t is the time trend, and $\varepsilon_{i,t}$ is the error term. Finally, $W^*_{i,t}$ is the tax-optimal gross wage with given total income $\Pi_{i,t} - R_{i,t} = W^g_{i,t} + D^g_{i,t}$. This is the variable of main interest in our analysis.[16] The parameter β denotes the average income-shifting effect on the actual gross wage income withdrawn from the firm.

The tax-optimal gross wage $W^*_{i,t}$ summarizes the effects that both the tax rate schedules of wages and dividends have on the actual realized gross wage. As we have the data actually used to tax the owners, we have all the information needed to define the tax-minimizing values $W^*_{i,t}$ and $D^*_{i,t}$ for every owner each year. The tax-optimal gross wage is calculated using tax register information on the owner's total gross income from the firm $(W^g_{i,t} + D^g_{i,t})$, net assets of the firm, gross earned income from other sources and

[15]In the data, the available controls for $X_{i,t}$ and C_i at the owner level are gender, age, other capital income, the ownership share of the firm and location dummies. On the firm level, the controls are turnover, number of employees, profits, total assets, and location and industry dummies.

[16]Fjaerli and Lund (2001) use a similar explanatory variable in their study.

the tax code and regulations for the year in question. In order to define $(W^*_{i,t}, D^*_{i,t})$ for each owner, we formulate a function that gives the tax-minimizing amount of wages and dividends for each possible total gross income level with respect to every combination of net assets and other earned income. Table A2 in the Appendix presents an illustrative example of the changes in tax optimal gross wages due to the tax reform of 2005.

As is well known in the microeconometric literature, estimating the causal effect of the tax code on the composition of realized income using equation (2.4.1) is difficult in practice. Many of the time-invariant variables that might affect income-shifting behavior are generally unobserved, which violates the exogeneity condition $cov(W^*_{i,t}, \varepsilon_{i,t}) = 0$. Therefore, we use panel data and the tax reform of 2005 to estimate the model. Taking first differences of equation (2.4.1) between t and $t+j$ gives us our estimable model

$$
\begin{aligned}
W^g_{i,t+j} - W^g_{i,t} &= (\alpha_{t+j} - \alpha_t) + \beta * (W^*_{i,t+j} - W^*_{i,t}) + \\
&\quad (X_{i,t+j} - X_{i,t}) + (\varepsilon_{i,t+j} - \varepsilon_{i,t}).
\end{aligned}
\tag{2.4.2}
$$

In this first-differences (FD) model, the time-invariant component C_i gets canceled out by definition. In contrast to the cross sectional one-year analysis in Fjaerli and Lund (2001), we focus on identifying the effect of the tax-optimal income component on the composition of income using exogenous individual variation in $W^*_{i,t}$ in time.

Our main interest is in the coefficient β, which expresses the average effect of a change in tax-optimal gross wages on the change in realized gross wages, conditional on given total gross income in t and $t+j$. The change in the tax-optimal gross wage $W^*_{i,t+j} - W^*_{i,t} = \triangle W^*_{i,t}$ captures all the changes in the individual tax code. In addition to changes in wage taxes, $\triangle W^*_{i,t}$ also captures changes in dividend and corporate taxation.

The testable hypotheses in the FD model are the following: If changes in the tax code explain the changes in the composition of income, β should be statistically significant and greater than zero. A one-to-one income-shifting response implies that $\beta = 1$. Also, adding control variables to the model should not affect the value of β, and the coefficients for the controls should not be statistically significant if the change in the tax code is the dominant factor behind the change in the division of income.

2.4.3. *Identification.* With regard to identifying parameter β, an important feature is that the tax reform of 2005 changed the income-shifting incentives differently among similar business owners. In other words, $\triangle W^*_{i,t} = W^*_{i,t+j} - W^*_{i,t}$ varies across otherwise similar individuals in the data. Owners with similar total gross income $(W^g_{i,t} + D^g_{i,t})$, other income, ownership share, firm total assets, profits and turnover but with different levels of firm net assets faced different changes in the marginal tax rates on dividends, and thus get different values of $\triangle W^*_{i,t}$. Owners with high level of net assets faced only modest changes in their marginal tax rates, whereas owners with low net assets faced larger tax incentives to rearrange their total gross income. Also, different levels of other earned income create variation in tax optimal gross wages, as income earned outside the firm affects the MTR on wages and progressively taxed dividends withdrawn from the firm. We assume that other earned income is exogenous.

A typical strategy in empirical tax research is to use marginal tax rates as explanatory variables when identifying behavioral effects of tax rate changes. However, using $\triangle W^*_{i,t}$ as a regressor instead of $\triangle(MTR_{W_{i,t}} - MTR_{D_{i,t}})$ helps to alleviate the issue of endogenous correlation between the income-shifting incentives and realized gross wages $W^g_{i,t}$. The optimal wage $W^*_{i,t}$ is not mechanically correlated with $W^g_{i,t}$ or $D^g_{i,t}$ at a given level of total gross income, whereas marginal or average tax rates themselves are. In most income tax systems, larger wages are associated with higher marginal tax rates and vice versa, causing these variables to be endogenously correlated in a FD model. However, realized gross wages do not affect the value of the tax-optimal gross wage, as $W^*_{i,t}$ is the same for *any* combination of $W^g_{i,t}$ and $D^g_{i,t}$ at a given level of $(W^g_{i,t} + D^g_{i,t})$.

Therefore, in the presence of exogenous tax rate variation, $\triangle W^*_{i,t}$ is exogenous in the FD model and does not necessarily require an instrumental variable. In contrast, marginal tax rates would need an instrument, and valid instruments for them are not widely available (see e.g. Saez et al. (2012)).

To identify β, we need to assume that in the absence of the reform, owners with a large positive $\triangle W^*_{i,t}$ do not change their $W^g_{i,t}$ differently than owners with smaller changes in $\triangle W^*_{i,t}$ (and vice versa). We have no explicit reason to assume that with given total income in t and $t+j$, the change in the realized gross wage $\triangle W^g_{i,t}$ depends on other factors than income-shifting incentives, *conditional* on individual and firm-level covariates. In the model, we control for other individual and firm-level variation in a rich way. In equation (2.4.2), $(X_{i,t+j} - X_{i,t})$ includes changes in the ownership share and other capital income on the owner's side, and changes in turnover, number of employees, profits and total assets on the firm side.

The empirical approach of using the tax-optimal income component as a measure for income-shifting is not solely linked to Finnish institutions or the dual income tax schedule. This approach generalizes to any case where there are two or more differently taxed tax bases available to the taxpayer. This also applies to different types of income which differ only with respect to tax deductions or allowances.

2.4.4. *Descriptive statistics*. One particular advantage of our empirical approach is that we can describe the extent of income-shifting behavior in a visually clear and convincing manner. After defining the tax-optimal combination of gross wages and gross dividends, we can compare the optimal gross wages to realized gross wages. Figure 2 presents the distribution of the difference between the tax-optimal gross wages and realized gross wages for the years 2002 and 2007. Tax-optimal behavior indicates that this difference would be equal to zero. In other words, $W^g_{i,t} - W^*_{i,t} = 0$ if the owner has optimized her wage 'perfectly' with respect to the tax code.

Figure 2 presents the distribution of $W^g_{i,t} - W^*_{i,t}$ around the tax-optimal point $W^g_{i,t} - W^*_{i,t} = 0$ in the range of +/- 10,000 € (in bins of 200 €). The Figure shows that income-shifting behavior is evident. There are clear spikes in the distribution at the level of 0 in both 2002 and 2007. Thus both before and after the reform a notable number of owners withdrew exactly the tax-optimal amount of wage income from the firm. This implies that the tax codes on both wages and dividends affect the total income composition of the owners, as there are no other explicit reasons for the owners to pay out exactly the tax-optimal amount of wages. In relative terms, over 40% of the owners in our sample optimized their wages perfectly in 2007. However, in 2002, we observe less complete wage optimization, as slightly under 15% of owners optimized their wages.

The monetary gains from income-shifting were smaller before 2005. This means that gains from optimizing the income composition are, on average, larger after the abolition of the single dividend tax system. This might explain the larger spike at the tax-optimal point after the reform in 2007. We further discuss the size of tax incentives and the costs of income-shifting in Section 2.5.

Figure 3 describes the relationship of the key variables in our study, the change in realized gross wages $\triangle W^g_{i,t} = W^g_{i,t+j} - W^g_{i,t}$ and the change in tax-optimal gross wages $\triangle W^*_{i,t} = W^*_{i,t+j} - W^*_{i,t}$ between the years 2002 and 2007. There is a clear positive relationship between the variables. On average, large $\triangle W^*_{i,t}$ are followed by similar $\triangle W^g_{i,t}$, which indicates that changes in the realized division of gross income are closely related to the changes in the tax code. Thus the owners who faced large changes in the tax-optimal income composition also changed their realized wages more than the owners who faced no or only small changes in tax incentives.

We fit a non-parametric Kernel estimate with a 95% confidence interval into Figure 3 to illustrate this effect and its statistical significance. Furthermore, the Figure illustrates

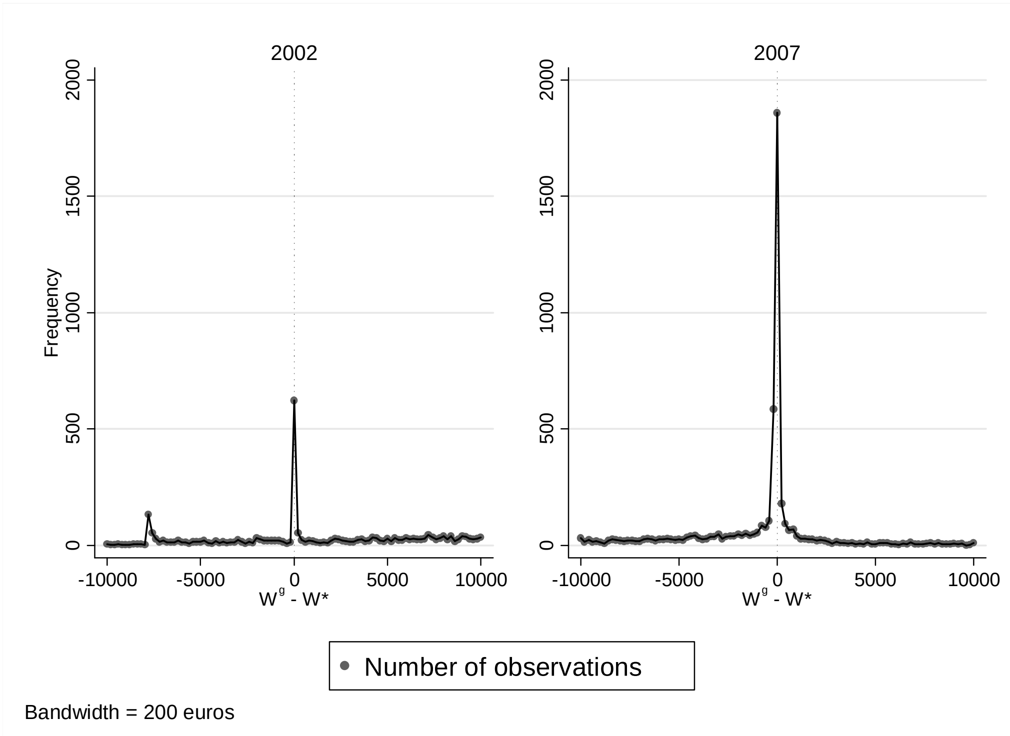


FIGURE 2. The distribution of the difference between realized gross wages and tax-optimal gross wages in 2002 (left) and 2007 (right)

that there is a considerable amount of variation in both realized and tax-optimal gross wages in the data.

Tables A3 and A4 in the Appendix present descriptive statistics for the key variables in our analysis. Table A3 presents the variables at the owner level, and Table A4 describes the characteristics at the firm level. Finally, Figure A1 in the Appendix presents the kernel density estimate distributions of wages and dividends received by the owners of privately held corporations both before (2002) and after (2007) the tax reform of 2005.

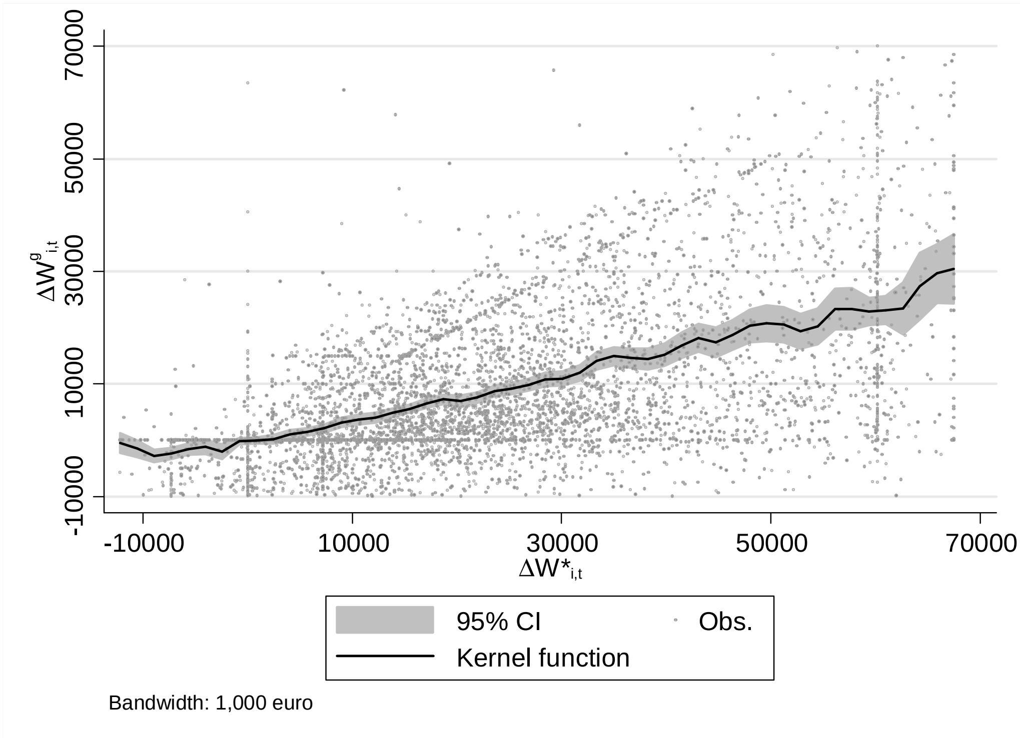


FIGURE 3. The effect of changes in tax-optimal gross wages $\triangle W^*_{i,t}$ on changes in realized gross wages $\triangle W^g_{i,t}$ between 2002 and 2007

2.5. Results

We estimate the first-differences equation (2.4.2) using a balanced panel data consisting of the years 2002, 2003, 2007 and 2008, and adding year dummies to the model. We estimate the equation in levels, as many observed and optimal wages and optimal dividends are zeros both before and after the reform. Therefore, for example, a logarithmic model would lose too much information.

The results are presented in Table 1. The first column shows the effect of a change in tax-optimal gross wages on a change in the realized gross wages without control variables. The second column estimates are derived using the full set of individual and firm-level controls.[17]

The owners of privately held corporations react to tax changes very actively. The tax schedule has a remarkable and statistically significant effect on the decision to divide income into wages and dividends. The coefficient for the optimal gross wage implies that a one euro change in the tax-optimal gross wage affects realized gross wages by 66 cents on average. The estimate differs statistically from 1, so the income-shifting response is not "perfect".

Adding control variables does not change the results. The coefficient for optimal gross wages with controls is very close to the coefficient without them, which supports the view that the tax schedule is the main factor affecting the income composition. Furthermore, adding controls does not affect the fit of the model. The R-squared statistic increases only by 0.01 compared to the model with $\triangle W^*_{i,t}$ as the only explanatory variable.

We also use a two-year difference model for the years 2002 and 2008 to estimate the longer-run average effect. These results are presented in Table A6 in the Appendix. When using the data for 2002 and 2008, the point estimate for income-shifting is approximately 0.68. This estimate is not statistically different from that using the panel data for all four years. This indicates that our results are robust and independent of the length of the difference.[18]

The coefficients of the control variables are mostly insignificant or very small, which again indicates that the changes in the tax system are the driving force behind the

[17]We also estimate the cross sectional model in equation (2.4.1) with a full set of control variables. The cross section OLS estimates for the years 2002, 2003, 2007 and 2008 are presented in Table A5 in the Appendix. The results show that the point estimates for the coefficients of tax-optimal gross wages (W^*) are between 0.90-1.05 and highly significant in every year. These results imply that income-shifting incentives and realized behavior seem to be highly correlated. Fjaerli and Lund (2001) get qualitatively similar results in their cross sectional analysis for Norway.

[18]The results are robust using all pairs of pre and post-reform years. Other results are available from the authors upon request.

decision on income composition. However, the ownership share appears to have a negative effect on realized gross wages. When ownership is concentrated, the owner has more power to make tax optimal decisions on income composition. In this case, increased ownership appears to open up a way to pay out more low-taxed dividends at the expense of wages (given the changes in the tax code). This result is also expected in the light of previous literature. Chetty and Saez (2010) find that tax-optimization is more active among corporate owners who own larger shares of the firm.

In addition, a change in the turnover of the firm has a positive and statistically significant effect on the difference in realized gross wages, although the size of the effect is very small. This can be interpreted as indicating that the growth of the firm (in the sense of turnover) has a small increasing effect on wage compensations, given the change in the tax code. All the other coefficients for firm-level controls are statistically insignificant, including the number of employees, profits and total assets. Therefore, changes in most of the firm-side variables have no significant effect on the division of income on average.

Figure 2 in Section 2.4.4 above gives indicative evidence which supports the hypothesis that costs and benefits matter in income-shifting behavior. The Figure shows that tax-optimal behavior is much more common after the reform of 2005. One explanation for this finding is that income-shifting became more profitable in monetary terms. After the reform, the introduced double taxation of dividends exceeding the imputed return increased the difference of dividend and wage tax rates in many cases, which also increased monetary gains from income-shifting.

In addition, we use a quantile regression approach to illustrate how the size of the incentive affects the response. In Figure 4, we plot the estimates at separate percentile points with the 95% confidence intervals using equation (2.4.2) with the full set of controls. The dashed line in the Figure denotes the average estimate.

VARIABLES	(1) ΔWage	(2) ΔWage
ΔW^*	0.662***	0.661***
	(0.007)	(0.013)
ΔOwnership share		-71.580**
		(33.259)
ΔTurnover^		0.129***
		(0.046)
ΔTotal assets^		0.200
		(0.206)
ΔProfits^		-0.167
		(0.176)
ΔEmployees		9.927
		(9.469)
ΔOther capital income^		-0.549
		(0.382)
Observations	17,237	17,237
R-squared	0.347	0.348

Notes: Owner-level clustered robust standard errors in parentheses.*** p<0.01, ** p<0.05. First-differences model estimated by OLS using balanced panel data for 2002, 2003, 2007 and 2008: the dependent variable is the difference in realized gross wages.

^ In 1,000 euros

Table 1. Estimation results

As can be seen from the Figure, the point estimates are larger at higher percentiles. At the 95th percentile point, the estimate is not statistically different from 1. In contrast, the estimates are smaller for those whose tax incentives were not affected as much by the tax reform. Thus it appears that the income-shifting response is positively correlated with the size of the incentive change.

To further study the heterogeneity of the income-shifting effect, we estimate the model by firm and owner-level characteristics. We categorize firms into four equally sized groups and estimate equation (2.4.2) separately for these groups. We use base-year (2002) turnover, total assets and the number of employees as continuous variables to study if there are differences in income-shifting responses with respect to the size of

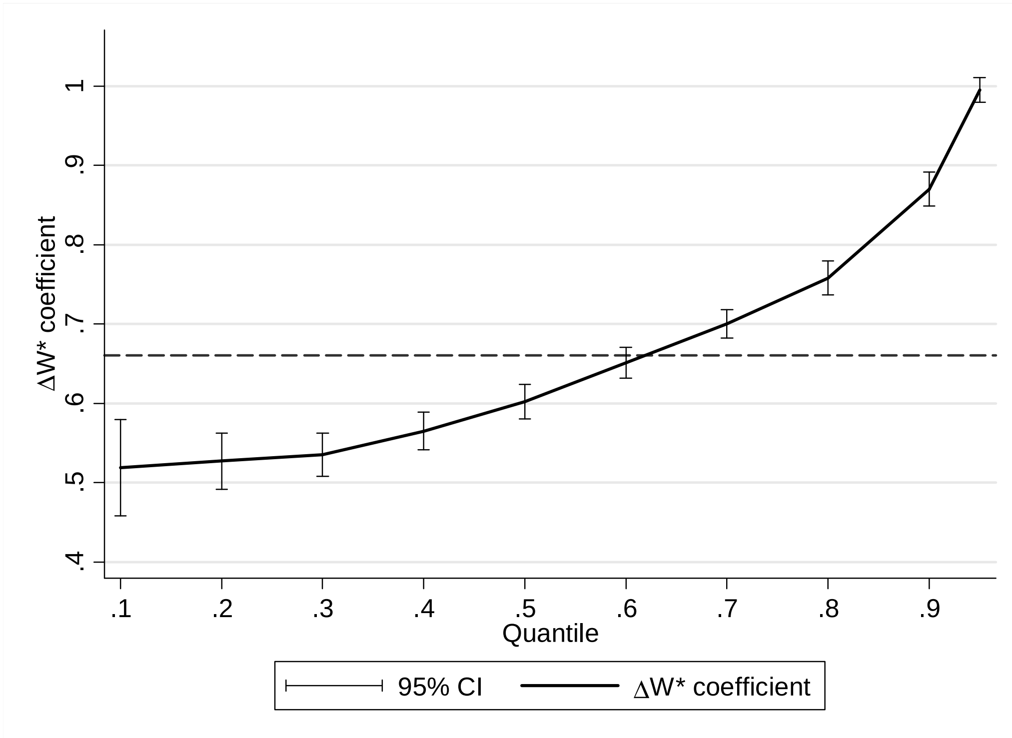


FIGURE 4. Quantile regression results

the firm. We also estimate the model by age and gender of the owner. In addition, we examine if there are differences in income-shifting activity between industries.

The results for different subgroups are presented in Table A7 in the Appendix. In general, the income-shifting responses are homogeneous between different groups. There are no significant differences in income-shifting activity between women and men, age groups or the size of the firm. Thus these results suggest that the average income-shifting response is not driven by certain types of owners or firms. However, some differences can be detected at the industry level. For example, the owners of firms in financing and agricultural industries shift income more actively than others.

There are some issues regarding the empirical setup that might affect the results. First, our data are limited to owners who receive dividends from their firms in each year.

This might bias the estimated average income-shifting effect among Finnish business owners. Also, the direction of the potential bias is somewhat unclear. The owners who do not pay any dividends might be more or less active in tax-motivated income-shifting compared to the owners who pay dividends. However, it is plausible that the owners not included in the data might be less active in income-shifting, especially before the reform of 2005 when there was in general larger tax incentives to pay dividends.

Second, our FD analysis uses balanced panel data for a relatively long time period (2002-2008). This implies that our estimating sample includes only owners who were successful enough to continue their business activity throughout this period. It might be that these owners are also more active in income-shifting. Thus this might cause an upward bias in our average estimate.

Third, our sample is limited to owners who own at least 50% of the firm alone or together with family members. It is presumable that these owners are more responsive to tax incentives than those who own less than 50%. The owners with more than 50% of the shares of the firm have more power to make tax-optimal decisions on profit distributions.

Finally, as mentioned in Section 2.2, pension and health insurance contributions might affect the income-shifting behavior. Insurance contributions are based on self-reported YEL income, which need not to coincide with the actual gross wage income of the owners in our estimating sample. However, wages and YEL income might be correlated among some owners. If insurance contributions are considered as taxes, this might decrease the incentives to increase wage payments as a response to dividend tax increase. This might create a downwards bias to our estimate, as we do not include insurance contributions based on YEL income as taxes when defining the tax-optimal wages.

2.6. Discussion

The main objective of this paper is to provide clear and intuitive evidence on the extent of income-shifting between tax bases among private business owners. We do this by relating the change in the actual income composition of the owner to the change in the tax-optimal income composition. In addition, we explore the heterogeneity of the income-shifting response among different owners and firms, and study how the size of tax incentives affect income-shifting behavior.

In many tax systems, business owners can minimize taxes by choosing an optimal combination of different income types as their personal compensation from the firm. The corporate and dividend tax reform of 2005 in Finland significantly changed the income-shifting incentives for many business owners. In the reform, the taxation of dividends tightened, which increased the incentives to pay wages as a form of personal compensation.

In the light of behavioral tax research, the reform had an appealing feature: the incentives to replace dividends with wages varied among approximately similar corporate owners. This variation in incentives together with extensive micro data, including information on both the owner and firm-level, enable us to credibly analyze the extent of income-shifting behavior.

We find strong and robust evidence that owners are active in income-shifting. Our main result shows that a one euro change in the tax-optimal gross wage results in a 66 cent change in realized gross wages on average. This indicates that the effect of the tax code on the composition of income is significant both statistically and economically. However, our estimate is statistically different from 1, which implies that not all of the owners behave according to a simple tax-minimizing model.

The income-shifting response appears to be homogeneous between different firms and owners. For example, the size of the firm does not affect the estimate. This implies that income-shifting behavior is not focused on certain types of firms and owners, and thus it cannot be explained by observable characteristics.

However, our results show that larger income-shifting incentives clearly increase the size of the income-shifting response. Quantile regression results indicate that responses are larger when incentives to shift income are larger. This (indirectly) implies that the monetary benefits from income-shifting affect the response.

These results suggest that a decrease in tax revenue caused by income-shifting can also be influenced by affecting the costs of tax optimization. At least to some extent, the costs and benefits of income-shifting can be affected by simply adjusting the tax regulations, and by decreasing the difference of the marginal tax rates on different tax bases. Finally, it is important to note that our analysis focuses solely on the income-shifting response. Tax rate changes might also have a significant effect on other behavioral margins, such as investments or entrepreneurial activity.

Bibliography

[1] Alstadsæter, A. and M. Jacob. (2012). Income shifting in Sweden – An empirical evaluation of the 3:12 rules. Report to the Expert Group on Public Economics (ESO) 2012:4, Swedish Ministry of Finance, Stockholm.

[2] Chetty, R. (2012). Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply. Econometrica, 80(3): 969–1018.

[3] Chetty, R. and E. Saez. (2010). Dividend and corporate taxation in an agency model of the firm. American Economic Journal: Economic Policy, 2(3): 1–31.

[4] Devereux, M., L. Liu and S. Loretz. (2014). The elasticity of corporate taxable income: New evidence from UK tax records. American Economic Journal: Economic Policy, 6(2): 19–53.

[5] Fjaerli, E. and D. Lund. (2001). The choice between owner's wage and dividends under the dual income tax. Finnish Economic Papers, 14: 104–119.

[6] Goolsbee, A. (2000). What happens when you tax the rich? Evidence from executive compensation. Journal of Political Economy, 108(2): 352–378.

[7] Gordon, R. and J. Slemrod. (2000). Are "real" responses to taxes simply income shifting between corporate and personal tax bases?, in Joel Slemrod (ed.), Does Atlas Shrug? The Economic Consequences of Taxing the Rich, Russell Sage Foundation and Harvard University Press, 2000: pp. 240–280.

[8] Kari, S., H. Karikallio and J. Pirttilä. (2008). Anticipating tax changes: Evidence from the Finnish corporate income tax reform of 2005. Fiscal Studies 29(2): 167–196.

[9] le Maire, D. and B. Schjerning. (2013). Tax bunching, income shifting and self-employment. Journal of Public Economics, 107: 1–18.

[10] Nielsen, S. B. and P. B. Sørensen. (1997). On the optimality of the Nordic system of dual income taxation. Journal of Public Economics, 63: 311–329.

[11] Pirttilä, J. and H. Selin. (2011). Income shifting within a dual income tax system. Evidence from the Finnish tax reform of 1993. Scandinavian Journal of Economics, 113(1): 120–144.

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

[13] Sivadasan, J. and J. Slemrod. (2008). Tax law changes, income-shifting and measured wage inequality: Evidence from India. Journal of Public Economics, 92: 2199–2224.

[14] Sørensen, P. B. (2005). Neutral taxation of shareholder income. International Tax and Public Finance, 12: 777–801.

Appendix

	MTR on wages		MTR on dividends (no net assets)		MTR on dividends (net assets 170k)		MTR on dividends (net assets 1,000k)	
Income	*2002*	*2007*	*2002*	*2007*	*2002*	*2007*	*2002*	*2007*
5,000	18.1	11.6	23.1	32.3	29.0	26.0	29.0	26.0
10,000	23.9	17.0	19.3	35.1	29.0	26.0	29.0	26.0
15,000	37.4	32.6	36.3	36.6	29.0	26.0	29.0	26.0
20,000	43.4	32.6	42.3	41.3	23.1	32.3	29.0	26.0
25,000	43.4	43.1	42.3	46.7	23.1	35.1	29.0	26.0
30,000	43.4	43.1	42.3	46.7	32.3	36.6	29.0	26.0
35,000	49.4	48.5	48.3	49.5	36.3	41.3	29.0	26.0
40,000	49.4	48.5	48.3	49.5	42.3	46.7	29.0	26.0
45,000	49.4	48.5	48.3	49.5	42.3	46.7	29.0	26.0
50,000	49.4	48.5	48.3	49.5	48.3	49.5	29.0	26.0
55,000	56.4	48.5	55.3	49.5	48.3	49.5	29.0	26.0
60,000	56.4	48.5	55.3	49.5	48.3	49.5	29.0	26.0
65,000	56.4	56.5	55.3	53.7	48.3	49.5	29.0	26.0
70,000	56.4	56.5	55.3	53.7	48.3	49.5	29.0	26.0
75,000	56.4	56.5	55.3	53.7	55.3	49.5	29.0	26.0
80,000	56.4	55.6	55.3	53.2	55.3	53.7	29.0	26.0
85,000	56.4	55.6	55.3	53.2	55.3	53.7	29.0	26.0
90,000	56.4	55.6	55.3	53.2	55.3	53.7	29.0	26.0
95,000	56.4	55.6	55.3	53.2	55.3	53.2	29.0	32.3
100,000	56.4	54.8	55.3	52.8	55.3	53.2	23.1	35.1

Notes:
MTR on wages is calculated with dividend income equal to zero, and vice versa. MTR on wages includes average municipal taxes, central government income taxes, automatic tax deductions and tax credits and average firm-level social security contributions (3%). MTR on wages does not include pension and health insurance contributions, as these are based on self-reported YEL income which is not determined by wage income (see Section 2). MTR on wages does not include deductions based on insurance contributions. MTR on dividends includes corporate taxes on withdrawn dividends (after 2005). MTR on dividends includes all automatic tax deductions and tax credits. MTR on progressively taxed dividends includes average municipal taxes and central government income taxes. Marginal tax rates are calculated using Stata and the Finnish JUTTA microsimulation model.

TABLE A1. Marginal tax rates (MTR) on wages and dividends with different levels of firm net assets, years 2002 and 2007 (in nominal euros)

Total gross income	Net assets	Tax optimal gross wage *2002*	Tax optimal gross wage *2003*	Tax optimal gross wage *2007*	Tax optimal gross wage *2008*
15,000	*10,000*	7,700	7,300	14,500	14,100
50,000	*10,000*	7,700	7,300	49,100	49,100
100,000	*10,000*	7,700	7,300	67,500	66,000
15,000	*100,000*	12,000	12,200	14,500	14,000
50,000	*100,000*	7,700	7,300	41,000	41,000
100,000	*100,000*	7,700	7,300	67,500	66,000
15,000	*500,000*	12,000	12,200	14,500	14,000
50,000	*500,000*	12,000	12,200	14,500	14,000
100,000	*500,000*	7,700	7,300	55,000	55,000

Notes:
The optimal gross wage levels are defined assuming that the owner owns 100% of the shares and that the owner has no earned income from other sources.

In general, earned income from other sources lowers the tax optimal gross wage, especially before the reform. For example, assume the owner has 2,500 € of other earned income with total gross income from the firm being 50,000 € and net assets 100,000 €. The tax optimal gross wage in 2003 is in this case 4,800 € (compared to 7,300 € without other earned income). However, with the same combination of total gross income, net assets and other earned income, the optimal gross wage does not change after the reform (41,000 € in both 2007 and 2008). This is due to the fact that after 2005 the tax rates for progressively taxed dividends increased sharply. After the reform, it is not in general optimal for the owner to replace wages with dividends after receiving a modest amount of other earned income.

Table A2. Tax-optimal gross wages before (2002, 2003) and after (2007, 2008) the 2005 tax reform with different levels of total gross income and net assets of the firm (in nominal euros)

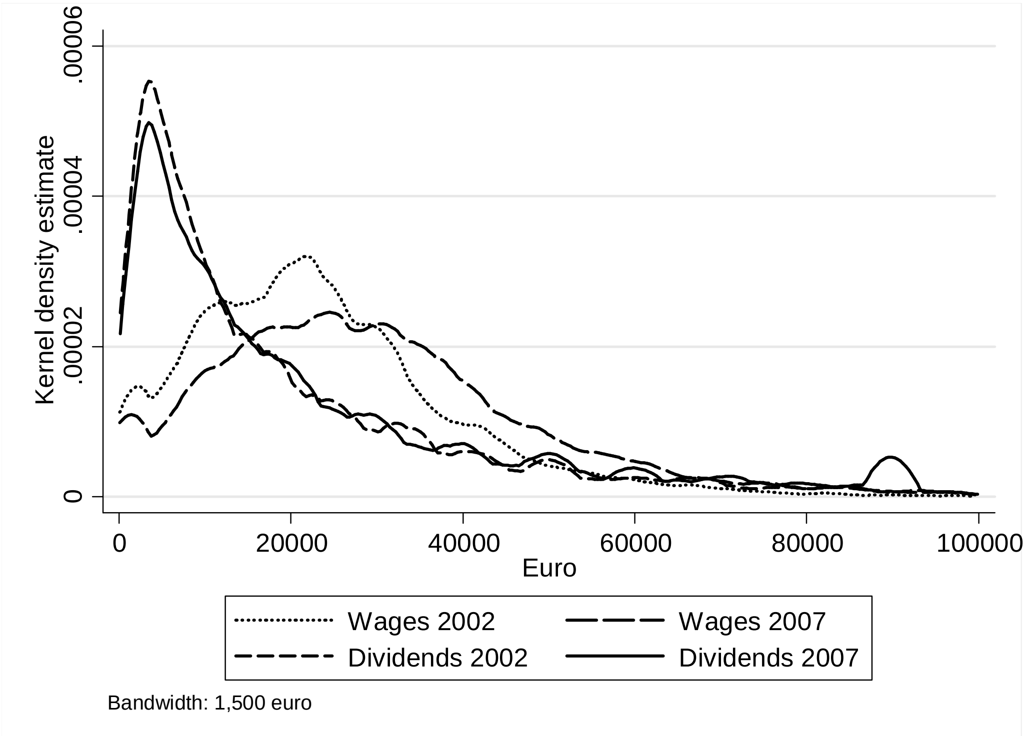


FIGURE A1. The distributions of wage and dividend income of the owners of privately held corporations in 2002 and 2007 (in nominal euros)

Year	Stat	Wages	Optimal wages	Dividends	Optimal dividends	Total income	Ownership share
2002	Mean	19,806	5,317	27,105	41,594	46,911	0.82
	Median	18,485	7,463	12,222	28,797	34,567	.93
	SD	16,986	3,499	82,510	84,965	85,066	0.23
	N	6,277	6,277	6,277	6,277	6,277	6,277
2003	Mean	19,244	4,794	32,744	47,194	51,988	0.84
	Median	17,223	7,011	15,000	31,783	36,996	.95
	SD	17,318	3,401	142,723	144,477	144,533	0.23
	N	6,277	6,277	6,277	6,277	6,277	6,277
2007	Mean	23,083	26,033	32,767	29,817	55,850	0.82
	Median	20,440	23,888	14,910	11,267	40,170	.99
	SD	22,443	19,416	99,552	100,123	102,931	0.22
	N	6,277	6,277	6,277	6,277	6,277	6,277
2008	Mean	23,980	26,233	35,487	33,234	59,468	0.82
	Median	20,880	23,739	15,400	12,680	42,300	.99
	SD	24,064	20,041	103,706	105,115	107,824	0.22
	N	6,277	6,277	6,277	6,277	6,277	6,277

TABLE A3. Descriptive statistics (2002, 2003, 2007 and 2008): Main owners (in nominal euros, estimation sample)

Year	Stat	Turnover	Employees	Total assets	Net assets
2002	Mean	782,450	10.35	400,805	285,155
	Median	227,617	4	141,598	100,222
	SD	4,092,140	32.98	2,174,166	1,669,665
	N	6,277	6,277	6,277	6,277
2003	Mean	946,741	10.27	529,807	381,950
	Median	289,713	4	192,240	114,693
	SD	3,982,281	30.64	2,375,763	5,233,616
	N	6,277	6,277	6,277	6,277
2007	Mean	1,082,630	10.60	723,319	448,007
	Median	321,193	4	253,792	152,155
	SD	3,155,168	36.14	2,985,295	2,378,661
	N	6,277	6,277	6,277	6,277
2008	Mean	1,152,018	10.63	811,968	516,807
	Median	329,951	4	272,411	168,326
	SD	3,329,805	36.25	3,452,935	2,791,899
	N	6,277	6,277	6,277	6,277

TABLE A4. Descriptive statistics (2002, 2003, 2007 and 2008): Firms (in nominal euros, estimation sample)

VARIABLES	(2002) Wage	(2003) Wage	(2007) Wage	(2008) Wage
W^*	1.050***	1.054***	0.904***	0.919***
	(0.075)	(0.071)	(0.014)	(0.015)
age	731.402***	796.057***	152.225	13.974
	(178.766)	(177.301)	(166.080)	(180.098)
age sq.	-8.102***	-9.032***	-1.295	0.104
	(1.912)	(1.852)	(1.650)	(1.771)
male	2,054.167***	1,887.503***	222.468	103.157
	(632.076)	(610.805)	(471.941)	(500.517)
ownership share	-5,615.921***	-6,330.395***	-3,311.677***	-1,888.356**
	(1,003.374)	(975.413)	(773.002)	(881.820)
turnover^	0.086	-0.033	0.190	0.351**
	(0.153)	(0.260)	(0.127)	(0.171)
total assets^	-0.227	0.988**	0.304*	0.338**
	(0.228)	(0.471)	(0.184)	(0.164)
profits^	8.841***	-0.275	-0.554*	-0.037
	(2.300)	(3.150)	(0.330)	(0.240)
employees	18.056	28.357	5.856	3.568
	(23.840)	(25.448)	(5.471)	(7.255)
capital income^	-1.132***	-11.207	0.610	
	(0.424)	(9.352)	(1.833)	
Constant	-4,342.653	-2,223.960	2,042.214	2,317.819
	(4,351.468)	(4,047.587)	(4,210.924)	(4,606.181)
Observations	5,160	5,611	6,244	6,237
R-squared	0.115	0.114	0.637	0.613

Notes: Owner-level clustered robust standard errors in parentheses.*** p<0.01, ** p<0.05, * p<0.1.
^ In 1,000 euros

TABLE A5. Cross-section results for the years 2002, 2003, 2007 and 2008

VARIABLES	(1) ΔW	(2) ΔW
ΔW^*	0.681***	0.680***
	(0.012)	(0.016)
ΔOwnership share		-9.120
		(52.054)
ΔTurnover^		0.130
		(0.083)
ΔTotal assets^		0.183
		(0.279)
ΔProfits^		-0.629
		(2.465)
ΔEmployees		-7.535
		(12.391)
ΔOther capital income^		-0.352
		(0.228)
Observations	5,613	5,613
R-squared	0.348	0.349

Notes: Owner-level clustered robust standard errors in parentheses.*** $p<0.01$.
^ In 1,000 euros

Table A6. Results for the years 2002 and 2008

VARIABLES	Turnover 0-25th p ΔW	Turnover 26-50th p ΔW	Turnover 51-75th p ΔW	Turnover 76-100th p ΔW	Employees 0-25th p ΔW	Employees 26-50th p ΔW
$\triangle W^*$	0.676***	0.597***	0.646***	0.613***	0.604***	0.626***
	(0.028)	(0.029)	(0.028)	(0.033)	(0.025)	(0.034)
Observations	1,528	1,529	1,529	1,529	2,009	1,387
R-squared	0.383	0.345	0.365	0.253	0.317	0.332

VARIABLES	Employees 51-75th p ΔW	Employees 76-100th p ΔW	Total assets 0-25th p ΔW	Total assets 26-50th p ΔW	Total assets 51-75th p ΔW	Total assets 76-100th p ΔW
$\triangle W^*$	0.606***	0.655***	0.738***	0.711***	0.640***	0.647***
	(0.027)	(0.033)	(0.027)	(0.024)	(0.024)	(0.033)
Observations	1,301	1,418	1,529	1,529	1,529	1,528
R-squared	0.377	0.302	0.359	0.417	0.380	0.262

VARIABLES	Age 0-25th p ΔW	Age 26-50th p ΔW	Age 51-75th p ΔW	Age 76-100th p ΔW	Male ΔW	Female ΔW
$\triangle W^*$	0.601***	0.628***	0.606***	0.583***	0.623***	0.590***
	(0.028)	(0.028)	(0.032)	(0.037)	(0.017)	(0.033)
Observations	1,597	1,587	1,623	1,308	5,247	868
R-squared	0.330	0.348	0.283	0.274	0.318	0.355

VARIABLES	Agriculture ΔW	Mining ΔW	Industry ΔW	Construction ΔW	Commerce ΔW	Hotels ΔW
$\triangle W^*$	0.836***	0.561***	0.692***	0.570***	0.600***	0.638***
	(0.108)	(0.081)	(0.048)	(0.035)	(0.030)	(0.092)
Observations	70	156	842	1,070	1,500	137
R-squared	0.537	0.394	0.335	0.308	0.322	0.430

VARIABLES	Logistics ΔW	Finance ΔW	Estate ΔW	Education ΔW	Health care ΔW	Other services ΔW
$\triangle W^*$	0.563***	0.964***	0.636***	0.693***	0.658***	0.579***
	(0.078)	(0.107)	(0.028)	(0.124)	(0.068)	(0.108)
Observations	462	63	1,433	48	208	125
R-squared	0.254	0.660	0.342	0.590	0.423	0.346

Note: Owner-level clustered robust standard errors in parentheses.*** p<0.01.

Table A7. Results for different subgroups, 2002-2008

CHAPTER 3

The Elasticity of Taxable Income and Income-shifting: What is "Real" and What is Not?[1]

ABSTRACT. Previous literature shows that income taxation significantly affects the behavior of high-income earners and business owners. However, it is still unclear how much of the response is due to changes in effort and other real economic activity, and how much is caused by tax avoidance and tax evasion. This distinction is important because it affects the welfare implications and policy recommendations. In this paper we distinguish between real responses and tax-motivated income-shifting between tax bases. We show how the explicit inclusion of income-shifting affects the welfare analysis of income taxation. In our empirical example we find that income-shifting accounts for over two thirds of the overall elasticity of taxable dividend income among Finnish business owners. The large income-shifting response significantly decreases the marginal excess burden compared to the standard model in which the overall elasticity defines the welfare loss. However, in addition to income-shifting, we find that dividend taxation significantly affects the real behavior of owners

Keywords: elasticity of taxable income, tax avoidance, income-shifting, real responses

JEL codes: H24; H25; H32

3.1. Introduction

Income taxes are known to generate significant behavioral effects among high-income earners and business owners (see a survey by Saez, Slemrod and Giertz 2012). However, the interpretation of the behavioral response is often difficult because business owners and high-income earners have many margins in which they can respond to taxes. In addition to real responses (labor supply, effort etc.), they have many opportunities to legally avoid or illegally evade taxes. Although previous research shows that tax

[1]This essay is joint work with Jarkko Harju (Government Institute for Economic Research VATT). Earlier versions of this paper are published in the Oxford University Centre for Business Taxation Working papers series, WP 13/13, November 2013, CESifo Working paper series, 4905, July 2014, and Government Institute for Economic Research Working paper series, 56, September 2014.

avoidance is a significant behavioral margin for these groups (see e.g. Slemrod and Gillitzer 2014), it is still unclear how much of the overall response is due to changes in real economic activity, and how much is due to avoidance.

Income-shifting is one of the most relevant tax avoidance channels for business owners and high-income earners (see e.g. Gordon and Slemrod 2000, Goolsbee 2000). Distinguishing between real responses and income-shifting between different tax bases is important because the nature of the response largely affects the welfare conclusions and policy recommendations (Slemrod 1995, Piketty, Saez and Stantcheva 2014). For example, real responses stemming from deeper behavioral parameters such as labor-leisure preferences are not under direct government control. In contrast, income-shifting can be governed more easily by re-designing the details of the tax system.

Traditionally, the elasticity of taxable income (ETI) quantifies the excess burden of the income tax (Feldstein 1995, 1999). However, income-shifting between tax bases is one of the most relevant issues that might overstate ETI as a measure of welfare losses. ETI with respect to its own marginal tax rate does not account for the fact that other tax bases might have positive tax rates. Thus income-shifting is not a full deadweight loss if the shifted income is also taxed (Saez 2004, Chetty 2009).

Our main contribution to the literature is to distinguish between real responses and income-shifting responses. We show how the explicit inclusion of income-shifting affects the welfare analysis of income taxation. Based on previous theoretical literature (e.g. Piketty et al. 2014), we build an empirically implementable model of ETI under income-shifting possibilities. Adding the difference of the net-of-tax rates on different tax bases to the standard ETI model enables identification of the average income-shifting elasticity and the average real elasticity. Furthermore, we demonstrate that different empirical specifications lead to different interpretations of the estimated parameters.

As an empirical example, we estimate both real responses and income-shifting responses for the owners of privately held corporations in Finland. This group faces large incentives and ample possibilities to shift income between different tax bases, which makes these owners a particularly suitable group for analyzing both income-shifting and real responses. In the Finnish dual income tax system, the owners of privately held corporations can withdraw income from their firm as a combination of wages and dividends, which are taxed with separate tax rate schedules and tax rules. There are only a few minor legal limitations on whether income is withdrawn as wages or dividends, and explicit tax rate differences induce clear incentives for tax-motivated income-shifting. To our knowledge, this paper is the first to explicitly estimate both real elasticity and income-shifting elasticity separately using a well-defined empirical model, and individual-level panel data and tax reforms.

We use extensive panel data of Finnish business owners. We link firm-level tax record information to the owner-level personal tax data, which is a novelty in the ETI literature. With this data set we are able to richly control for firm-level effects on the personal income trends of the owners. The comprehensive data along with the dividend tax reform of 2005 in Finland creates an interesting opportunity to study the role of both income-shifting and real income creation.

Our results show that the income-shifting responses are highly significant both statistically and economically. Over two thirds of the overall ETI among Finnish business owners is due to income-shifting. However, income-shifting does not appear to be the whole story, as we also find positive real elasticity estimates for dividends. In addition, real responses are present even when analyzing broader firm-level income components, such as turnover and profits. These are less subject to tax avoidance than wages and dividends withdrawn from the firm. The tax elasticities of these firm-level income components are also rarely analyzed in public finance literature.

Our results highlight that welfare evaluations based on standard ETI estimates might be misleading for individuals with income-shifting possibilities. For example, the marginal excess burden of dividend taxes decreases from 0.9 to 0.4 when we account for the fact that the shifted income is also taxed. The large income-shifting effect also affects policy recommendations. Even though dividends appear to be very responsive as a whole, dividend taxes do not induce substantial distortions in the real economy among the owners of privately held corporations in Finland.

The ETI literature began to expand after the pioneering studies by Lindsey (1987) and Feldstein (1995). Feldstein (1995) estimates the taxable income elasticity to be large in the US, ranging from 1-3 depending on the income group. Many subsequent studies focus on improving the identification of the ETI model. Along with the refinements, the estimates have decreased markedly. A wide range of studies report average elasticity estimates from 0 to 0.6. For example, the widely cited Gruber and Saez (2002) study finds ETI of 0.2 for mid-income earners, and 0.6 for high-income earners. A review of earlier empirical results is presented in the recent survey by Saez et al. (2012).

Recently, the literature has identified the behavioral response using income distributions around the discontinuous kink points of the marginal income tax rate schedule. Saez (2010) shows that excess bunching around kink points is proportional to the local ETI at the kink. Many studies show that the excess mass around kink points is larger for self-employed individuals (see Saez 2010, Chetty et al. 2011 and Bastani and Selin 2014). This indicates that the self-employed respond actively to income tax rates, and have more opportunities to adjust their behavior to them. As an additional analysis, we also estimate the local tax responsiveness of Finnish business owners using the bunching method. We find that business owners bunch actively at the dividend income tax rate kink point, which supports our main results.

Previous studies from different countries indicate that income-shifting between tax bases is substantial for high-income earners and business owners. For example, Gordon and Slemrod (2000) show evidence of active income-shifting between corporate and personal tax bases in the US. Devereux et al. (2014) show that income-shifting between corporate and personal tax bases is also active in the UK. In addition, Sivadasan and Slemrod (2008) find significant income-shifting responses for partners in partnership firms in India, and Romanov (2006) finds income-shifting between personal and corporate tax bases among high-income self-employed professionals in Israel.

Piketty et al. (2014) formulate a theoretical framework for analyzing tax avoidance effects as a part of the ETI of high-income earners. By distinguishing between different forms of behavioral responses (tax avoidance, real responses and bargaining channels), they study the implications of optimal income taxation at the upper end of the income distribution. They also provide empirical cross-country evidence which indicates that both the real and avoidance responses are small while bargaining effects dominate.

In the Nordic countries, le Maire and Schjerning (2013) derive a dynamic extension to the bunching method, and show that over half of the bunching effect among Danish entrepreneurs is due to intertemporal income-shifting. This result suggests that the excess burden calculated by using the baseline bunching method overestimates the welfare effect. By using panel data methods, Kleven and Schultz (2013) estimate the cross-tax elasticities of taxable earned income and taxable capital income components within the Danish tax system. In general, they find small substitutability between earned income and capital income, which, however, supports the view that income-shifting effects exist. In addition, Alstadsæter and Jacob (2012) show that income-shifting is active among Swedish corporate owners.

In Finland, Harju and Matikka (2012) show that absent any real effects, income-shifting between tax bases is active among the main owners of privately held corporations in Finland. Pirttilä and Selin (2011) show evidence of responses to the dual income tax reform in Finland in 1993. They report that entrepreneurs and business owners increased their relative share of capital income when capital income tax rates were decreased.

This paper is organized as follows: Section 3.2 presents the theoretical model. Section 3.3 presents our empirical model. Section 3.4 describes the Finnish income tax system and recent tax reforms. Section 3.5 discusses identification issues, introduces the data and presents the descriptive statistics. Section 3.6 presents the results, and Section 3.7 presents the robustness checks. Section 3.8 discusses the main findings and welfare implications.

3.2. Theoretical model

3.2.1. Taxable income model. Following Piketty et al. (2014), we assume a quasi-linear utility function of the form $u_i(c, z) = c - h_i(z)$, where c is consumption, z is taxable income, and $h_i(z)$ denotes the cost of effort to produce income for individual i. The cost function is assumed to be convex and increasing in z. Utility is maximized under the budget constraint $c = z(1 - \tau) + R$, where $(1 - \tau)$ is the net-of-tax rate (one minus the marginal tax rate) on a linear segment of a non-linear tax rate schedule. R denotes virtual income.

Optimization of the utility function with respect to the budget constraint results in individuals producing taxable income up to the point where $h_i'(z) = (1 - \tau)$. Thus in the absence of income effects, individual taxable income supply is a function of $(1 - \tau)$.

Next, consider a marginal change in $(1 - \tau)$. The elasticity of taxable income (ETI) can be written as

$$e_z = \frac{(1-\tau)}{z} \frac{dz}{d(1-\tau)} \tag{3.2.1}$$

where e_z is the average ETI. In addition to changes in labor supply, e_z also covers changes in, for example, work effort and productivity. In addition, the average ETI covers tax avoidance and tax evasion.

The intuition behind this Feldstein (1999) framework is that all behavioral responses affect the excess burden of income taxation. Individuals increase z until its marginal cost equals the tax rate, and thus the overall inefficiency can be summarized with ETI. This requires that the marginal cost of effort, the marginal cost of tax avoidance and the marginal cost of tax evasion etc. all equal the net-of-tax rate. In other words, $h_i'(z) = (1-\tau)$ no matter how z is adjusted, and thus estimating e_z is all we need for welfare analysis.

3.2.2. Taxable income and income-shifting. The standard ETI in equation (3.2.1) implicitly includes income-shifting from or to another tax base due to a change in $(1-\tau)$. However, among other previous papers that discuss the implications of tax avoidance in ETI analysis (e.g. Saez et al. 2012, and Chetty 2009), we argue that more precise modeling of income-shifting is needed. With regard to welfare analysis, this is essential if ETI is analyzed among individuals who have easy access to differently taxed tax bases.

Income-shifting can be very difficult for the average wage earner due to the lack of opportunities to alter her income composition. However, it can have a major impact for individuals who indeed have these possibilities. In general, high-income earners and business owners have more ways to affect the composition of their personal income. For example, in many countries business owners have opportunities to report part of their personal taxable income as corporate profits, or vice versa.

In particular, income-shifting opportunities are apparent within a dual income tax system where capital income and wage income are taxed differently with separate tax rules and regulations. In the Finnish dual income tax system, the most prominent income-shifting incentives lie between the wage and dividend income of the owners of privately held corporations. We discuss the Finnish system in more detail in Section 3.4.

We present a static taxable income model with income-shifting opportunities. Our model is similar to the elasticity of taxable corporate income model by Devereux et al. (2014), and the Piketty et al. (2014) model with tax avoidance in the top income bracket.[2]

We assume that there are two personal tax bases available, taxable wages z_W and taxable dividends z_D. We denote the total taxable income of the owner by $z_y = z_W + z_D$. In many tax systems, business owners have many different channels to withdraw income from their firm. Our model generalizes to any two differently taxed tax bases in which an individual can legally report income.

Wages are taxed at a tax rate τ_W, and dividends are taxed at τ_D. It is possible for the owner to shift income (at a cost) between the two types of income. The owner has an incentive to shift income from one tax base to another if the tax rate schedules differ from each other. Intuitively, income-shifting behavior describes the extent of changing the composition of income due to differences in τ_W and τ_D, while keeping the level of total taxable income constant.

For simplicity, let us assume for now that $t_W > t_D$. This is usually the case in most dual income tax systems. We assume that both tax rates are exogenous.[3]

The budget constraint can be written as

$$c = (1 - \tau_W)(1 - \alpha)z_y + (1 - \tau_D)\alpha z_y \tag{3.2.2}$$

where $0 \leq \alpha \leq 1$, and $(1 - \alpha)z_y = z_W$ is taxable wages denoted as a share of total taxable income. Similarly, $\alpha z_y = z_D$ is taxable dividends.

The utility function of an owner i is

$$u_i(c, z_y, \alpha) = c - \theta_i(z_y) - \phi_i(\alpha) \tag{3.2.3}$$

[2]Other previous papers also consider tax avoidance and income-shifting within the ETI framework, e.g. Saez (2004) and Chetty (2009).

[3]Income-shifting from wages to dividends produces more total net income for the owner if $\tau_W > \tau_D$. Naturally, the opposite direction for income-shifting holds if $\tau_W < \tau_D$. If the tax rates are equal, we are back to the case of one common tax base.

where $\theta_i(z_y)$ is the cost of effort to produce total taxable income, and $\phi_i(\alpha)$ is the cost of income-shifting between wages and dividends, i.e. changing the composition of total taxable income. We assume that both cost functions are convex and increasing in z_y and α, respectively.[4]

The owner chooses z_y and α to maximize utility, which gives the following first-order conditions:

$$(1-\tau_W)(1-\alpha)+(1-\tau_D)\alpha=\theta_i^{'}(z_y) \tag{3.2.4}$$

and

$$(\tau_W-\tau_D)z_y=\phi_i^{'}(\alpha) \tag{3.2.5}$$

Equation (3.2.4) implies that total taxable income is an increasing function of the net-of-tax rates. Thus when α is assumed to be fixed, both tax rates affect the total taxable income, as we have two separately taxed tax bases with no income-shifting possibilities. Equation (3.2.5) implies that when keeping the amount of total taxable income (z_y) fixed, income-shifting is an increasing function of the tax rate difference. Thus the difference between the tax rates, $(\tau_W-\tau_D)$, determines the amount of income shifted from one tax base to another.

Next, we derive elasticities separately for both tax bases. For both z_W and z_D, we assume no changes in the tax rate of the other tax base. Following Piketty et al. (2014)

[4]Alternatively, we could assume that both real wages and real dividends have separate convex cost functions that reflect real wage and real dividend income based on labor supply and effort, and the actual return on invested capital. This type of model gives qualitatively similar results as the model with one cost function for all income. In addition, we could assume that ϕ_i is fixed for each i, and does not depend on α. However, this does not change the welfare conclusions based on estimated average elasticities (see Section 3.2.3).

and Devereux et al. (2014), the average net-of-tax rate elasticity for z_W is

$$
\begin{aligned}
e_{z_W} &= \frac{(1-\tau_W)}{z_W}\frac{\partial z_W}{\partial(1-\tau_W)} \\
&= \frac{(1-\tau_W)}{(1-\alpha)z_y}\frac{\partial z_y}{\partial(1-\tau_W)}(1-\alpha) + \frac{(1-\tau_W)}{(1-\alpha)z_y}\frac{\partial(1-\alpha)}{\partial(1-\tau_W)}z_y \\
&= e_W - e_{(1-\alpha)}
\end{aligned}
\tag{3.2.6}
$$

where $e_W = dz_y/z_y * (1-\tau_W)/d(1-\tau_W)$, and $e_{(1-\alpha)} = d(1-\alpha)/(1-\alpha) * (1-\tau_W)/d((1-\tau_D)-(1-\tau_W))$.

Equation (3.2.6) implies that we can distinguish the income-shifting effect from the overall behavioral response e_{z_W}. The income-shifting elasticity $e_{(1-\alpha)}$ measures how the wage tax base reacts to changes in the difference of the net-of-tax rates. We refer to the other component e_W as the real elasticity. It denotes how total income changes as the wage tax rate changes. We discuss the practical limitations of interpreting e_W as an actual real effect in Section 3.3.

Similarly, we can express the average ETI of dividend income as

$$
\begin{aligned}
e_{z_d} &= \frac{(1-\tau_D)}{z_D}\frac{\partial z_D}{\partial(1-\tau_D)} \\
&= \frac{(1-\tau_D)}{\alpha z_y}\frac{\partial z_y}{\partial(1-\tau_D)}\alpha + \frac{(1-\tau_D)}{\alpha z_y}\frac{\partial \alpha}{\partial(1-\tau_D)}z_y \\
&= e_D + e_\alpha
\end{aligned}
\tag{3.2.7}
$$

where $e_D = dz_y/z_y * (1-\tau_D)/d(1-\tau_D)$ is the real dividend elasticity, and $e_\alpha = d\alpha/\alpha * (1-\tau_D)/d((1-\tau_D)-(1-\tau_W))$ is the income-shifting elasticity for dividends.

In summary, equations (3.2.6) and (3.2.7) differ from (3.2.1) as they take income-shifting explicitly into account. As noted in Piketty et al. (2014), z_W and z_D are more responsive to changes in their own net-of-tax rates than without income-shifting possibilities (or with arbitrarily large costs for income-shifting). However, if income-shifting is not important in practice, $e_{(1-\alpha)}$ and e_α should be small or insignificant.

3.2.3. Welfare implications. Next, we compare the marginal excess burden in the traditional ETI model with a model that explicitly includes income-shifting. Our model for the marginal deadweight loss follows the one presented in Chetty (2009). We approximate the marginal excess burden by comparing behavioral responses caused by a tax rate change to a benchmark case which ignores behavioral responses. The same follows from assuming that the tax revenue collected from wage and dividend taxes is returned to the owner as a lump-sum transfer.

We use the following welfare function

$$w = \{(1-\tau_W)(1-\alpha)z_y + (1-\tau_D)\alpha z_y - \theta_i(z_y) - \phi_i(\alpha)\} + (1-\alpha)z_y\tau_W + \alpha z_y\tau_D \tag{3.2.8}$$

where individual utility is presented in curly brackets, and tax revenue collected by the government is denoted as the sum of the tax revenue from both tax bases. We again assume that $\tau_W > \tau_D$.

Let us first consider the standard deadweight loss analysis of wage taxation. Conceptually this refers to the simplified case where $\alpha = 0$. The same analysis can be carried out for dividends, but for the sake of brevity we only show the equations for taxable wage income.

Consider a marginal increase in the wage tax rate, $d\tau_W$. As the owner is assumed to optimize her utility, we can use the envelope theorem and denote that the tax increase has only a first-order effect on individual utility. The first-order effects of the owner's utility and the tax revenue of the government cancel each other out. Thus we can write the excess burden as

$$\frac{dw}{d\tau_W} = \tau_W \frac{\partial z_y}{\partial \tau_W} = z_y \frac{\tau_W}{(1-\tau_W)} e_{z_W} \tag{3.2.9}$$

where e_{z_W} denotes the overall elasticity of the wage tax base with respect to $(1-\tau_W)$. In equation (3.2.9), e_{z_W} refers to the standard ETI in the Feldstein (1999) framework. Intuitively, in equation (3.2.9), average ETI defines the scope of the marginal excess burden of the income tax.

Next, consider a more general case where the owner can shift part of her taxable wage income to the dividend tax base. This refers to the case where $0 \leq \alpha \leq 1$, and owners can adjust α. The deadweight loss is expressed as

$$
\begin{aligned}
\frac{dw}{d\tau_W} &= \frac{\partial z_y}{\partial \tau_W}\left((1-\alpha)\tau_W + \alpha\tau_D\right) + z_y \frac{\partial(1-\alpha)}{\partial \tau_W}(\tau_W - \tau_D) \\
&= z_y \left[\frac{(1-\alpha)\tau_W + \alpha\tau_D}{(1-\tau_W)} e_W + (1-\alpha)\frac{(\tau_W - \tau_D)}{(1-\tau_W)} e_{(1-\alpha)}\right]
\end{aligned}
\tag{3.2.10}
$$

where e_W denotes the average real elasticity, and $e_{(1-\alpha)}$ is the average income-shifting elasticity.

The key difference between equations (3.2.9) and (3.2.10) is the income-shifting response. Assume that we observe an overall decrease in taxable wage income due to an increase in the wage tax rate, $e_{z_W} > 0$. Assume further that part of this response comes in the form of income-shifting, $e_{(1-\alpha)} > 0$, and part of the response is due to changes in real economic behavior, $e_W > 0$. If we ignore the income-shifting response and use the standard equation (3.2.9) to assess the marginal excess burden, the welfare effect is approximated to be too large when $0 < \tau_D < \tau_W < 1$ and $0 \leq \alpha \leq 1$.

The size of the marginal excess burden in equation (3.2.10) depends on the following factors: (1) the size of the income-shifting elasticity ($e_{(1-\alpha)}$), (2) the size of the real elasticity (e_W), (3) the difference of the net-of-tax rates ($\tau_W - \tau_D$), and (4) the initial size of the tax bases $(1-\alpha)$. Intuitively, a large $e_{(1-\alpha)}$ relative to e_W implies that a large fraction of the overall response is due to income-shifting, with different efficiency implications. For a given $e_{(1-\alpha)}$, a small $(\tau_W - \tau_D)$ implies that income-shifting has only a small effect on efficiency, and vice versa. In addition, the relative size of the tax bases further scales the significance of the income-shifting response. To summarize, if there are large incentives for income-shifting, equation (3.2.10) highlights that it is important to estimate elasticities for both the real component and the income-shifting component in order to more accurately analyze the welfare loss of income taxes.

Equation (3.2.10) shows that income-shifting and real responses have different welfare consequences even within the standard excess burden framework when the shifted income is also taxed. Applying the envelope theorem in the welfare model above implicitly assumes that individuals optimize such that the marginal costs equal the associated net-of-tax rates (see the first-order conditions (3.2.4) and (3.2.5) above). Thus equations (3.2.9) and (3.2.10) hold if individuals optimize as in the standard Feldstein (1999) framework. However, it is possible that these standard assumptions do not hold in practice, especially when considering the welfare effects of income-shifting.

Chetty (2009) shows that the Feldstein (1999) formula for the deadweight loss does not hold if the marginal *social* cost of income-shifting does not equal the tax rate. For example, as noted by Chetty (2009), a notable share of the costs related to income-shifting might be payments to tax consultants, who usually report at least part of this original cost as their own taxable income. Thus the costs might include transfers between different agents in the economy, and this fiscal externality is not taken into account in the standard framework. In the extreme, if income-shifting inflicts no real social costs, the marginal excess burden reduces to the real effect of taxation, denoted by the first term on the right-hand side of equation (3.2.10). We further discuss this and the empirical magnitude of the deadweight loss among Finnish business owners in Section 3.8.

3.3. Empirical methodology

3.3.1. Empirical ETI model. A usual approach to estimate ETI with individual-level panel data and tax reforms is to use a difference-in-differences approach and a first-differences estimator.[5] This method allows time-invariant unobserved individual characteristics that affect income growth to be canceled out. This is appealing as these characteristics (for example, innate ability) are potentially correlated with the progressive marginal tax rate.

Following Saez et al. (2012), the standard empirical ETI equation can be characterized as

$$\triangle ln(z)_{t,i} = e_z \triangle ln(1-\tau)_{t,i} + \triangle ln(\eta)_{t,i} + \triangle ln(\varepsilon)_{t,i} \tag{3.3.1}$$

where t is a subscript for time and i denotes the individual, and $\triangle$ denotes the difference between time $t+k$ and t. z denotes taxable income, $(1-\tau)$ is the net-of-tax rate, and e_z is the (average) elasticity of taxable income. η denotes potential income, i.e. income without taxes, and ε is the error term, including the transitory income component.

There are many issues that need to be taken into account when defining the actual empirically implementable version of equation (3.3.1). First, the net-of-tax rate and transitory income shocks are mechanically correlated within a progressive tax system, as a positive income shock results in a lower net-of-tax rate. This means that a valid instrument for the net-of-tax rate is required in order to have a causal interpretation for e_z.

Non-tax-related changes in potential income also need to be taken into account. In other words, differential income growth trends for different types of individuals need to

[5] For a more detailed discussion on empirical ETI estimation, including cross-sectional models, see Saez et al. (2012). We discuss the local estimation of ETI using distributions of taxable income and bunching around the kink points of the tax rate schedule in Section 3.7.1.

be controlled for. The usual approach is to add a matrix of individual characteristics in base year t to the estimable equation

3.3.2. ETI and income-shifting. In this subsection we specify ways of distinguishing income-shifting from the overall response using micro-level panel data and tax reforms. For now we assume that valid net-of-tax rate instruments exist, and that we can perfectly control for other individual characteristics that affect the growth of taxable income. These issues will be discussed in detail in Section 3.5.

First, in order to identify different elasticity components, we need to have differential variation in marginal tax rates among otherwise similar individuals. This variation is needed for all relevant tax bases. In the case of Finnish business owners, we need variation in both the wage and dividend tax rates.

By utilizing exogenous variation in the net-of-tax rates on both wages and dividends, we can write the estimable version of the elasticity of taxable wage income as

$$\begin{aligned} \triangle ln(z_W)_{t,i} =& e_W \triangle ln(1-\tau_W)_{t,i} - e_{(1-\alpha)} \triangle (ln(1-\tau_D) - ln(1-\tau_W))_{t,i} \\ & + \triangle ln(\eta_W)_{t,i} + \triangle ln(\varepsilon)_{t,i} \end{aligned} \tag{3.3.2}$$

where $(1-\tau_W)$ is the net-of-tax rate for taxable wage income, and $(1-\tau_D)$ is the net-of-tax rate for dividend income.

Equation (3.3.2) includes the responsiveness of taxable wage income with respect to income-shifting incentives, namely $\triangle(ln(1-\tau_D)-ln(1-\tau_W))_{t,i}$. Regressing $\triangle ln(z_W)_{t,i}$ with both $\triangle ln(1-\tau_W)_{t,i}$ and $\triangle(ln(1-\tau_D)-ln(1-\tau_W))_{t,i}$ enables us to estimate separately both the real elasticity e_W and the income-shifting component $e_{(1-\alpha)}$, along with the associated standard errors. A similar model can also be written for dividend income. For the sake of brevity, we only cover the wage income model in this section.

The income-shifting effect can also be estimated by simply adding $\triangle ln(1-\tau_D)_{t,i}$ to the standard ETI model for taxable wages. After adding dividend tax rates, we get the following expression

$$\triangle ln(z_W)_{t,i} = e_{z_W}\triangle ln(1-\tau_W)_{t,i} - e_{(1-\alpha)}\triangle ln(1-\tau_D)_{t,i} + \triangle ln(\eta_W)_{t,i} + \triangle ln(\varepsilon)_{t,i} \tag{3.3.3}$$

Importantly, adding $\triangle ln(1-\tau_D)_{t,i}$ to the standard ETI model does not change the interpretation of the baseline ETI parameter e_{z_W}, which captures *both* real responses and the income-shifting effect. If income-shifting behavior is significant, the estimated sum of these elasticity components might not be very informative. Therefore, the standard ETI model alone or even conditional on the net-of-tax rates in other tax bases might be misleading when assessing the welfare consequences of income taxation. Nevertheless, in terms of identifying the income-shifting response, both equations (3.3.2) and (3.3.3) define the same income-shifting elasticity parameter $e_{(1-\alpha)}$.

Another possibility to separate out the income-shifting response is to study the elasticity of total taxable income $z_y = z_W + z_D$ with respect to the net-of-tax rate on wages. In the earlier literature this type of income has in many cases been referred to as broad income (see e.g. Gruber and Saez 2002). The model for the total taxable income can be written as

$$\triangle ln(z_y)_{t,i} = e_W\triangle ln(1-\tau_W)_{t,i} + \triangle ln(\eta_y)_{t,i} + \triangle ln(\varepsilon)_{t,i} \tag{3.3.4}$$

The elasticity coefficient in equation (3.3.4) only includes the real response component, as any income-shifting is canceled out by definition. In other words, if an increase in the wage tax rate induces only a pure income-shifting effect, total taxable income remains unchanged. Thus regressing $\triangle ln(z_y)_{t,i}$ with $\triangle ln(1-\tau_W)_{t,i}$ allows us to identify the real elasticity component, which can then be compared with the taxable wage income elasticity in order to outline the relevance of income-shifting behavior.[6]

To summarize, how well we can estimate both e_W and $e_{(1-\alpha)}$ depends on the data we have. We can outline real responses with total income data including all the relevant tax bases. In order to analyze both e_W and $e_{(1-\alpha)}$, we need information on the wage tax base separately. Also, differential and independent variation in both marginal tax rates is necessary for identifying $e_{(1-\alpha)}$ separately. In addition, in order to analyze underlying differences in the responsiveness of the wage and dividend tax bases, we need to estimate the dividend tax base elasticities as well. As we meet all the conditions mentioned here with our data set, we can study how different specifications affect the estimates in our empirical example. These conditions are also applicable to other tax systems in other countries.

We are particularly interested in estimating the effect of tax rates on real behavior, in relation to income-shifting effects. In other words, we want to exclude any other forms of tax avoidance (or evasion) when estimating e_W. Therefore we use gross wage and gross dividend income subject to taxation as dependent variables when estimating the models. These income measures do not include potential changes in deduction behavior. Instead, in the Finnish context, taxable income is defined as gross income subject to taxation minus deductions and exemptions. Thus taxable income also takes into account changes in deduction behavior, which presumably also include changes in tax avoidance activity.

Despite using changes in gross income as the left-hand side variable, interpreting e_W as a true real response includes an implicit assumption that income-shifting and tax

[6] If we include changes in both net-of-tax rates in equation (3.3.4), we cannot identify both real responses and income-shifting responses separately. The estimated coefficients would be a mixture of both real and income-shifting components of wage and dividend taxes, as before in the standard ETI model (equation (3.3.1)). Therefore, we examine real responses by estimating total income regressions separately for both net-of-tax rates, which allows us to identify the real response component for both net-of-tax rates.

deductions are the only possible margins of tax avoidance. However, other possibilities to avoid taxes might be included in the estimated real response. For example, if tax rates increase, owners could increase their consumption within the firm (e.g. in the form of more office amenities). Also, owners might increase fringe benefits, which are in many cases not fully included in gross income subject to taxation. Finally, owners might illegally evade taxes, for example through intentional underreporting of income.

One of the most common examples of other tax avoidance channels affecting e_W is intertemporal (or dynamic) income-shifting. For example, dividend income can be rather easily shifted across periods using retained earnings. In the Finnish context, Kari et al. (2008) show evidence that Finnish corporations anticipated the 2005 dividend tax increase by increasing dividend payments just before the reform. Anticipation was feasible as the content of the reform was published already in late 2003. Mostly due to this anticipation possibility, our baseline empirical analysis utilizes a longer time period of 2002-2007.

Furthermore, in order to assess the real component in a more diverse manner, we estimate the net-of-tax rate responses for more broadly defined income components at the firm level. One example of these is net profits before wages. We define net profits as turnover plus other income of the firm minus all costs except wages. Compared to wages and dividends withdrawn from the firm, this type of income is not as easily manipulated using various tax avoidance activities. In addition to wages and dividends, net profits also includes retained earnings. Intuitively, changes in net profits due to changes in net-of-tax rates reflect the real effort of the owner.

In addition, we estimate net-of-tax rate elasticities for the turnover of the firm. Turnover measures the overall sales revenue of the firm, which also reflects real effort and productivity.[7] We futher discuss the details of these estimations in Section 3.6.

[7] Harju and Kosonen (2013) study the tax responsiveness of turnover among the owners of unincorporated firms in Finland. They find small real responses for this group.

3.3.3. Estimable equation. We estimate different variations of the following equation using a two-stage least squares estimator:

$$\triangle lnTI_{t,i} = \alpha_0 + e\triangle ln(1-\tau^p)_{t,i} + \alpha_1 f(lnTI)_{t,i} + \alpha_2 B_{t,i} + \alpha_3 F_{t,i} + \triangle\varepsilon_{t,i} \tag{3.3.5}$$

In equation (3.3.5), $\triangle lnTI_{t,i}$ is the log change in income between t and $t+k$. The income concept varies across different specifications. First, we analyze the responsiveness of gross wage income and gross dividend income subject to taxation with respect to own net-of-tax rates and income-shifting incentives. Thus in these cases we set $TI_{t,i} = (z_W)_{t,i}$ for wages, and $TI_{t,i} = (z_D)_{t,i}$ for dividends. In addition to the two separate tax bases, we also regress the change in total income $TI_{t,i} = (z_W + z_D)_{t,i}$ with changes in the instrumented net-of-tax rates. We also estimate alternative models for real responses, where we regress broader firm-level income components, namely turnover and net profits, with changes in the instrumented net-of-tax rates.

$\triangle ln(1-\tau^p)_{t,i}$ is the instrumented change in the log net-of-tax rate (we discuss the instruments in detail in Section 3.5). Thus e is the coefficient of interest, the average elasticity with respect to the net-of-tax rate. When studying income-shifting responses, we add the difference of the log net-of-tax rates of wages and dividends into the estimable equation.

Following Gruber and Saez (2002), we add a 10-piece base-year income spline $f(lnTI)_{t,i}$ to the model. Base-year income controls for unobserved heterogeneity in income growth. We also control for observed individual effects with available background variables in the tax return data. Matrix $B_{t,i}$ includes age, age squared, ownership share of the firm, and county and gender of the owner. In addition, firm-level data allow us to control for firm-level effects. The firm-level controls $F_{t,i}$ include total assets, turnover, profits, industry, number of employees and county of the firm.

In our baseline model, we analyze a single difference between 2002 and 2007. As is common procedure in the literature, we focus on the owners at the intensive margin whose firms are their primary source of income. We limit the analysis to observations where base-year total income (wages + dividends) is above 25,000 €. In addition, individuals whose absolute change in total income between 2002 and 2007 is above 50,000 € are dropped from the sample in order to avoid unnecessarily high influence by outlier observations. We perform several robustness checks on these sample restrictions in Section 3.7.1.

All the estimates are weighted by the total income of the owner. When considering the welfare consequences of income taxation, income-weighted uncompensated average ETI is the parameter of main interest (see Gruber and Saez 2002). However, as in Gruber and Saez (2002), we censor the weights at 200,000 € in order to avoid giving unreasonably large weight to a few very high-income individuals in the data.

In addition to first-differences estimation, we also use the distributions of z_W and z_D and the kink points in the marginal tax rate schedules to estimate ETI locally. We discuss this bunching estimation in more detail in Section 3.7.1.

3.4. Finnish income tax system and recent tax reforms

In our empirical example we analyze the owners of privately held corporations in Finland. Privately held corporations are defined as corporations that are not listed on a public stock exchange (cf. public or listed corporations). In the Finnish tax system, dividends from listed and privately owned corporations are taxed at different tax rates and under different tax regulations. Also, the taxation of privately held corporations is different from that of other types of private businesses (sole proprietors and partnerships). Furthermore, we focus on tax reforms that occurred in 2002-2007, as we use the same time period in our baseline analysis.

Since 1993, Finland has applied a dual income tax system where earned income (wages, pensions, fringe benefits etc.) and capital income (interest income, capital gains, dividends from listed corporations etc.) are taxed separately. Earned income is taxed on a progressive tax rate schedule (0-56% in 2007), whereas the capital income tax rate is flat (28% in 2007).[8] A typical feature of the Nordic dual income tax system is that the top marginal tax rate on earned income is much higher than the flat tax rate on capital income. The lower flat tax rate on capital income was justified on various grounds, for example broadening the tax base and decreasing the scope for tax arbitrage, and increased global capital mobility.

Within the Finnish dual income tax system, wage income and dividend income from privately held corporations are taxed with separate tax rules and tax rates. In general, owners of these firms can freely choose the income composition of wages and dividends, and income-shifting between these tax bases is legitimate. There are only a few minor legal limitations on whether income is withdrawn as wages or dividends from a privately held corporation in Finland. Therefore, for example, reporting more dividend income at the expense of wages induces no fines or penalties.

However, wages cannot be paid without a work contribution for the firm, or else wages may be considered as veiled distribution of profits. In addition, dividends can be paid only if the firm has distributable assets. These include, for example, accumulated profits and non-tied equity. Nevertheless, in contrast to wages and dividends, other alternatives for withdrawing income from the firm are restricted. These include, for example, shareholder loans and share repurchases. In summary, as income-shifting responses between wages and dividends are largely unrestricted among the owners of

[8] As a whole, the Finnish income tax system follows the principle of individual taxation. The income of a spouse or other family members does not affect the marginal income tax rate of an individual. However, some tax deductions and social security benefits depend on the total income of the household.

privately held corporations in Finland, analyzing this group provides an intuitive empirical example to study the significance of both income-shifting and real behavioral incentives.

3.4.1. Dividend taxation and the dividend tax reform of 2005. The Finnish dual income tax system includes specific rules for the dividend taxation of the owners of privately held businesses. Dividends are categorized into two parts according to the net assets (assets-liabilities) of the firm:[9]

- The amount of dividends corresponding to an imputed 9% return on the net assets of the firm are subject to a flat tax rate. The imputed rate of return on net assets is set by the government, and it is the same for all owners.
- Any dividends exceeding the imputed return are taxed progressively.

For example, with net assets of 400,000 € and the imputed rate of return set at 9%, the maximum amount of dividends taxed at the flat tax rate is 36,000 € (0.09*400,000=36,000). In other words, any dividends from the firm up to 36,000 € are taxed at the flat tax rate. Any dividend income *exceeding* this amount is taxed according to the progressive tax rate schedule.

Furthermore, the taxation of dividends exceeding the imputed return is not similar to that of wage income. Dividends are subject to corporate taxes whereas wages are not. Also, some tax deductions are only allowed on wage income, whereas dividends are not subject to firm-level social security contributions.

Before the 2005 tax reform, a full imputation system of corporate taxes was in place. Within the full imputation system, corporate taxes paid on distributed dividends were credited back to the shareholder, which led to the effective single taxation of dividend

[9]The net assets of the firm are calculated using the asset and debt values in the year before. The individual net asset share of the owner is calculated based on the ownership share of the firm. Also, there are some individual adjustments to the net assets. For example, if the owner or her family members live in a dwelling which is owned by the firm, the value of this dwelling is not included in net assets when calculating the imputed return.

income. Thus, before 2005, both flat-tax and progressively taxed dividends were only subject to individual-level taxes.

The reform of 2005 changed the principle of dividend taxation, as the full imputation system was replaced by double taxation of dividends. After the reform, all dividends became subject to the 26% corporate tax rate. In addition to corporate taxes, all dividends are in principle also taxed at the individual level. However, only 70% of dividend income was subject to individual taxation after the reform.

Nevertheless, single taxation of dividend income was partly retained. Dividends below the imputed 9% return on net assets and below 90,000 € remained single-taxed (at the flat 26% corporate tax rate). Therefore, only dividends exceeding the imputed return or 90,000€ were subject to the double taxation rule.

The reform changed the marginal tax rates (MTR) differently for different types of owners. In general, changes in the MTR on dividends depend both on the amount of dividends and the net assets of the firm. Table 1 presents the main changes in the MTR on dividends for different types of owners.

Effective marginal tax rates on dividends (D)		
	Before (2002)	After (2007)
Type (I): D $\leq$ Imputed return and D $\leq$ 90,000€	29%	26%
Type (II): D $\leq$ Imputed return and D $>$ 90,000€	29%	40.5%
Type (III): D $>$ Imputed return		
min	0%	26%
max	55%	54%

TABLE 1. Effective marginal tax rates on dividends before (2002) and after (2007) the reform of 2005 for different types of owners

The first type of owners (Type (I)) are those who have dividend income below the 9% imputed return on net assets and below 90,000 €. For these owners the effective flat tax rate on dividends decreased from 29% to 26%. Before the reform, dividends below the imputed return were not subject to the corporate tax rate, and were taxed only at the flat personal capital income tax rate of 29%. After the reform, these dividends are only subject to the 26% corporate tax rate, and are not taxed at all in individual taxation.

Type (II) owners are those who have dividend income below the imputed return on net assets and above 90,000 €. Before the reform, these dividends were taxed at the flat capital income tax rate. After the reform, 70% of dividends above 90,000 € are regarded as taxable capital income in personal taxation, in addition to the flat corporate tax rate of 26%. This results in an effective flat tax rate of 40.5% for these dividends after the reform, compared to 29% before the reform.

Type (III) owners are those who have dividend income above the imputed return on net assets. Before the reform, these dividends were only taxed as personal earned income, subject to a progressive tax rate schedule (0-55%). After the reform, 70% of dividends above the imputed return are regarded as taxable earned income, in addition to the flat corporate tax rate of 26%. Therefore, the reform significantly increased the MTR on small dividends exceeding the imputed return, but the changes in the MTR were small for large dividends above the imputed return on net assets.

In summary, owners with larger net assets were more likely to be faced with a decrease in their dividend tax rate. In contrast, owners with smaller net assets were more likely to face an increase in their marginal dividend tax rate. Therefore, otherwise similar owners who differ only in the net assets of the firm were faced with different changes in the marginal tax rate on dividends. This implies that the change in the MTR on dividends is not directly related to the amount of dividend income, which alleviates the usual identification issues in the literature. We discuss these in more detail in Section 3.5.1.

Figure A1 in the Appendix presents the effective marginal tax rates on dividends in 2002 and 2007 with two levels of net assets, 0 and 250,000 € (approximately the average net assets in the estimation sample before the reform). The Figure shows that most of the MTR increases occur on low and middle dividend income exceeding the imputed return. Also, the Figure shows the 3 percentage point drop in the flat tax rate on dividends below the imputed return and 90,000 €.

Plans to abolish the single taxation of dividend income were widely discussed already in 2002, and more formal propositions were published by the Finnish Government in late 2003. Thus it was possible for owners to anticipate the shift to the double taxation system. Therefore, the years right before and right after the reform are not suitable for empirical analysis that aims at identifying longer-run behavioral parameters.[10] Also, in 2005, special transition rules were applied which alleviated the partial double taxation of dividends. Thus, in our empirical analysis, we exclude the years 2003-2006 from the regression, and focus on the medium-run effect of the tax reform using a single difference of 2002-2007.

Finally, the main motivation behind the reform of 2005 was not the economic and fiscal conditions in Finland. The pre-reform full imputation credit was granted only to domestic shareholders whose firms operate in Finland. This violated European Union rules on the equal tax treatment of all EU citizens. Thus Finnish legislators were more or less forced to change the tax system towards more unified treatment of domestic and international shareholders. Therefore, the tax reform of 2005 can be considered exogenous from the point of view of domestic shareholders.

3.4.2. Wage income taxation and variation in wage tax rates. In Finland, there are three levels of wage income taxes: central government (or state-level) income taxes, municipal income taxes and mandatory social security contributions. The central government income tax rate schedule is progressive, whereas municipal tax rates and social security contributions are proportional by nature. Municipal tax rates vary

[10] Kari, Karikallio and Pirttilä (2008) provide empirical evidence of anticipation effects before the tax reform of 2005 for privately held corporations.

between different municipalities.[11] Social security contributions include, for example, unemployment insurance payments.

During 2002-2007, there was a general decline in central government income tax rates throughout the income distribution. Marginal tax rates decreased almost every year in most income classes within central government taxation. In contrast, municipal tax rates have changed differently in different municipalities. On average, every fifth municipality changed its tax rate in each year. Yearly municipal tax rate changes vary from -1 to +1.5 percentage points, which accounts for roughly 1-10% changes in the overall net-of-tax rate. On average, the municipal tax rate increased from 17.8% in 2002 to 18.5% in 2007.

Because of different tax rate changes in different municipalities, the marginal wage tax rates of owners have changed differently. Also, municipal tax rate changes are determined only by the municipality of residence, not by the income level of an individual owner. Furthermore, since municipal taxation is residence-based, the marginal wage tax rate of the owner is not determined by the municipality in which the firm is physically located or registered.

Figure A2 in the Appendix describes the MTR on wage income. The left-hand side of the figure shows that average marginal wage tax rates decreased throughout the income distribution in 2002-2007. The right-hand side of Figure A2 shows the actual marginal tax rates calculated using our data set for the year 2007, highlighting the fact that individuals with the same income level face different marginal tax rates due to municipal-level tax rate differences. In addition, owners with the same income level face different changes in the MTR on wages due to different changes in municipal tax

[11] There are 336 municipalities in Finland (in 2012). Each democratically elected municipal council decides on the municipal tax rate on an annual basis. Municipalities can choose their tax rates freely. However, certain legislative municipal-level duties need to be financed mainly by municipal taxes (e.g. basic health care and primary education).

rates. This improves the identification of the elasticity parameter, since a notable part of the variation in wage tax rates is not directly based on taxable income.

We do not include mandatory pension and health insurance contributions as a direct tax on wages in this study. The insurance contributions of the owners of privately held corporations are not levied on actual wage income if the ownership share is above 50%, and the shareholder holds an executive position in the firm. These owners are termed YEL owners. YEL owners report a self-selected YEL income from which the insurance payments are accumulated. The reported YEL income can be above or below the actual wages paid without implications or sanctions.[12]

In contrast, insurance contributions are based on actual wage income from the firm for owners whose ownership share is less than 50% (similarly as in the case of paid workers with no ownership share). These owners are termed TEL owners. Thus for TEL owners, insurance contributions increase or decrease based on the wage income withdrawn from the firm. However, it is not clear whether insurance contributions are fully regarded as taxes, since owners directly benefit from them in terms of future pensions or in case of illness. Nevertheless, it is plausible that insurance payments levied on actual wage income decrease the incentives to pay wages to TEL owners, compared to YEL owners whose insurance contributions do not depend on wages withdrawn from the firm. We discuss how we apply this variation in our empirical analysis in Section 3.5.1.[13]

3.4.3. Tax incentives for income-shifting. The Finnish tax system creates noticeable income-shifting incentives for the owners of privately held corporations. As

[12]However, there are both lower and upper limits for YEL income, which are both also independent of actual wage income. Nevertheless, as insurance payments determine pensions after retirement as well as many income-bound social benefits, YEL owners have incentives to report a realistic YEL income that reflects the actual earnings potential.

[13]There were no relevant changes in TEL or YEL insurance payment rates in the time period we study. The average rate for TEL payments is 21.1% in both 2002 and 2007, and 20.8% in 2007 and 21.1% in 2002 for YEL payments. Insurance contributions are fully deductible from taxable income.

the tax rate schedules for wages and dividends differ from one another, owners can minimize income taxes by choosing an optimal combination of wages and dividends as their personal compensation from the firm. Harju and Matikka (2012) show that the owners of privately held corporations are active in minimizing tax payments through income-shifting between wages and dividends.

The 2005 tax reform affected the income-shifting incentives of many owners. In the light of our analysis, it is significant that the reform changed the income-shifting incentives differently among otherwise similar owners. Owners with dividends below the imputed return on net assets and 90,000 € faced only modest changes in their income-shifting incentives. For these owners, the flat tax rate on dividends decreased by 3 percentage points, inducing a small change in incentives to increase dividend compensation at the expense of wages. In contrast, owners with dividends exceeding the imputed rate of return on net assets faced an increase in dividend tax rates, as these dividends became double-taxed. For many of these owners, the MTR on dividends became larger than the MTR on wages, inducing notable changes in the incentives to shift income between wages and dividends.

Table A1 in the Appendix presents the marginal tax rates on wages and dividends at different levels of firm net assets. The table highlights that owners with different net assets have different MTRs on dividends, and faced different changes in marginal tax rates and income-shifting incentives from the 2005 tax reform.[14]

3.5. Identification and data

3.5.1. Net-of-tax rate instruments. In a progressive income tax rate schedule, the marginal tax rate increases as taxable income increases. Therefore, an increase in taxable income mechanically decreases the net-of-tax rate, causing the tax rate variable

[14] Harju and Matikka (2012) provide a more detailed discussion on income-shifting incentives within the Finnish dual income tax system.

to be endogenous in the empirical model. Thus a valid instrumental variable for the net-of-tax rate is required.

A common strategy in the ETI literature is to simulate predicted (or synthetic) tax rates and use them as instruments for the net-of-tax rate (NTR) (see Gruber and Saez 2002). The basic structure of the predicted NTR variable is the following: Take pre-reform income in base-year t, and use it to predict the net-of-tax rates for $t + k$ by using the post-reform tax legislation in $t + k$. The predicted tax rate instrument is then defined as the difference between the actual NTR in t and the NTR calculated with income in t and the tax law for $t + k$. Intuitively, the predicted NTR instrument describes the change in tax liability caused by changes in tax legislation, ignoring any behavioral effects via taxable income responses.

In this study we use the predicted NTR instrument with a few modifications. First, we need to address the development of net assets when defining the net-of-tax rate instrument for dividends. Net assets are a key factor determining the marginal tax rate on dividends (see Section 3.4). As shown in Tables A2 and A3 in the Appendix, average net assets increase in time both in the whole data set and our estimation sample. Thus we need an estimate for net assets in $t + k$ when defining the NTR instrument for dividends. Otherwise we would be predicting the effect of the dividend tax rate change incorrectly for a large number of owners, and the NTR instrument would be too weak.

We predict net assets after the reform for each owner using exogenous pre-reform characteristics in 2000-2003. We use the same exogenous individual and firm-level variables as in the baseline ETI regression. These variables include, for example, age, age squared, gender, turnover, total assets and industry and location dummies. Intuitively, counterfactual net assets take into account the development of net assets not related to the tax reform of 2005. The R-squared statistic for the net assets prediction using OLS is 0.73.

Second, predicted NTR instruments are better predictors of exogenous tax rate variation within a single tax base and a single tax bracket of the progressive tax rate schedule. Intuitively, predicted NTR instruments perform better for changes in income that are relatively close to the original income level in the base period. However, available income-shifting opportunities might cause substantial changes in taxable income, as income-shifting may lead to "jumping" across tax brackets. Therefore, the predicted net-of-tax rate instruments might be too weak if income-shifting is active.

Therefore, we might need additional instruments to more robustly estimate ETI for individuals with income-shifting possibilities. The purpose of additional instruments is to capture the incentives to change the composition of income, which are not necessarily taken into account when using only the predicted NTR approach.

We use the pension insurance status of the owner as an additional instrument in the wage regression. Pension insurance status is defined based on the ownership share of the firm, and the official working status of the owner of the firm. Individuals who work in their own firm in an executive position and own 50% or more of the firm alone or together with immediate family members are termed YEL owners. They can choose the amount of reported YEL income on which mandatory insurance payments are levied. In contrast, individuals who own less than 50% of the firm pay pension insurance payments based on the actual wages paid from the firm. These owners are called TEL owners. YEL/TEL status cannot be freely chosen. Owners satisfying the YEL conditions in a given year cannot change their status to TEL owners, or vice versa.

We assume that TEL owners who cannot choose the level of insurance payments would not increase their wage compensation after the reform as much as YEL owners, for whom wages do not legally affect the level of the insurance contribution. In other words, the YEL/TEL status affects the incentives to shift income from dividends to wages because of insurance payments, which are not captured by the predicted net-of-tax rate instrument. The exclusion restriction behind this instrument is that the

YEL/TEL status is not itself correlated with transitory income shocks, conditional on various observed individual and firm-level characteristics.

An essential issue in identifying ETI is the variation in marginal tax rates. With both wages and dividends, changes in the marginal tax rates vary across the income distribution. This is important because non-tax-related changes in income are potentially problematic when identifying the elasticity parameters (see Saez et al. 2012). If the shape of the income distribution varies independently of tax reforms, the analysis of behavioral responses to tax rate changes might be biased if this variation cannot be properly taken into account. Non-tax-related changes in the income distribution are especially problematic if the variation in MTR is focused only on a certain part of the income distribution, for example the tax rate cuts or increases in the top income bracket. The fact that both dividend and wage tax rate variation occurs in all income classes alleviates the potential problems associated with these issues.

Furthermore, the fact that changes in both the MTR on wages and dividends are not direct functions of income improves the exogeneity of the instrument. As discussed in the recent ETI literature, there is no proof that the predicted NTR instrument is exogenous in all cases (Blomquist and Selin 2010, Weber 2014). It is unlikely that the instrument is correlated similarly with both parts of the transitory income component ($\varepsilon_{t+k,i} - \varepsilon_{t,i}$) if the NTR is a direct function of income (even conditional on base-year income splines and other controls).[15] This is lessen of an issue in our empirical example, as changes in the MTR on both dividends and wages also depend on net assets and the municipality of residence, respectively.

[15]Blomquist and Selin (2010) use income in the middle year of the difference ($(t + k + t)/2$) as the base period when imputing the predicted tax rates for both $t + k$ and t. Weber (2014) proposes an instrument which exploits years before k as the base period. Both of these approaches aim at reducing the covariance between the instrument and the transitory error component. However, both of these strategies provide more or less weaker instruments than the Gruber and Saez (2002) approach, which might also bias the estimated parameter of interest.

3.5.2. Data. Our data are from the Finnish Tax Administration, and include information on the financial statements and tax records of Finnish businesses. The data include tax record information on both the firm and its main owner from the year 2000 onward. Another unique characteristic of the data is that they include all Finnish businesses (all public and private corporations, partnerships, sole proprietors etc.). In this study we focus on the main owners of privately held corporations.

The data set contains all important information for our analysis, for example wages and dividends paid to the owner by the firm, and income earned by the owner from other sources. These, together with other tax record information, enable us to define the marginal tax rates for both the relevant personal tax bases, wages and dividends. By linking the owner-level and the firm-level data together, we can control for various individual and firm-level effects in the empirical estimation. This type of detailed business owner data are rarely used in ETI analysis.

For this study we construct a balanced panel data set for the years 2002-2007. The main owner data include only those individuals who received positive dividends from the firm during a tax year. Tables A2 and A3 in Appendix describe the data and the key variables we use from both 2002 and 2007. Table A2 shows the statistics for the whole data, and Table A3 for our baseline estimation sample.

3.5.3. Descriptive statistics. Figure 1 describes the means of wage, dividend and total income (wages+dividends) from 2000 to 2009 for all owners of privately held corporations. The figure shows that mean wages, mean dividends and mean total taxable income all increased from 2000 to 2009. Importantly, the figure indicates that the share of wage income relative to total income has increased from 2005 onwards. This suggests that the tax reform of 2005 and the abolition of the single taxation of dividends affected the composition of total income, which gives us the first preliminary evidence that income-shifting might be significant. However, based on Figure 1, it remains unclear whether the tax rate changes also induced real responses.

In addition, Figure 1 shows that mean dividends and mean total income increased in the year before the reform, and decreased right after the implementation of the reform of 2005. Plans to abolish the single taxation system were publicly discussed already in 2002, and more formal propositions were published by the Finnish Government in late 2003. The early discussions also included a proposal to increase tax rates on dividends below the imputed return, which, however, eventually remained single-taxed after the reform. Nevertheless, from the point of view of business owners, various proposals and active public discussion increased the uncertainty surrounding the details of the 2005 tax reform.[16]

For owners with progressively taxed dividends, there were large incentives to withdraw extra dividend income from the firm before the double taxation came into effect in 2005. However, many owners with flat-tax dividends also increased dividend payouts before 2005. This was presumably caused partly by the uncertainty about the actual implementation of the new dividend tax system. Nevertheless, most owners faced at least some intertemporal incentives to increase dividend payments before the reform, as larger dividend payments became subject to higher marginal tax rates after the reform.

Next, we describe how tax incentives affect income withdrawn from the firm. Figure 2 shows the proportional changes in dividends (left-hand side) and total income (right-hand side) for two groups: those who faced a modest dividend tax decrease or no changes in the dividend tax rate, and those who faced a dividend tax increase. These groups are defined based on the predicted tax rate change due to the reform of 2005, calculated using the income information in 2002.[17] In the Figure, the light-gray dashed lines

[16]Kari et al. (2008) discuss the various proposals and their details more thoroughly.

[17]The predicted tax rate changes are defined similarly as the net-of-tax rate instruments (see Section 3.5.1 above). Owners with no changes in tax incentives include those with a change below 5% in the net-of-tax rate on dividends (in either direction). Owners with a dividend tax increase include owners

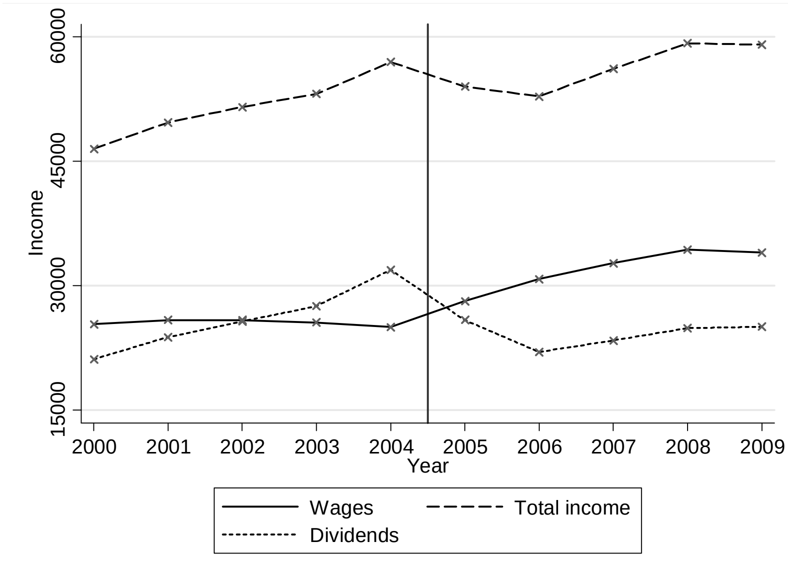


FIGURE 1. Means of wage, dividend and total income in 2000-2009 (in 2000 euros)

denote the anticipation period (2003-2005). The black dotted lines within 2003-2005 characterize the development of dividends and total income ignoring the anticipation period.

Figure 2 highlights the following issues: First, when comparing dividends in the pre-reform (2000-2002) and post-reform (2006-2009) periods, we can see that dividends decreased among owners who faced a predicted dividend tax increase. In comparison, dividends increased among owners with no tax rate changes or a dividend tax decrease. This indicates that owners responded to the dividend tax reform according to changes in tax incentives. Second, dividends increased in both groups in a similar fashion before 2003, which indicates that there were no significant differences in pre-reform income trends. Third, both groups anticipated the reform by notably increasing dividends

with a positive change above 5%, and owners with a dividend tax decrease include owners with a negative change above 5% in the predicted net-of-tax rate.

before the reform. This behavior is consistent with short-run intertemporal income-shifting incentives.

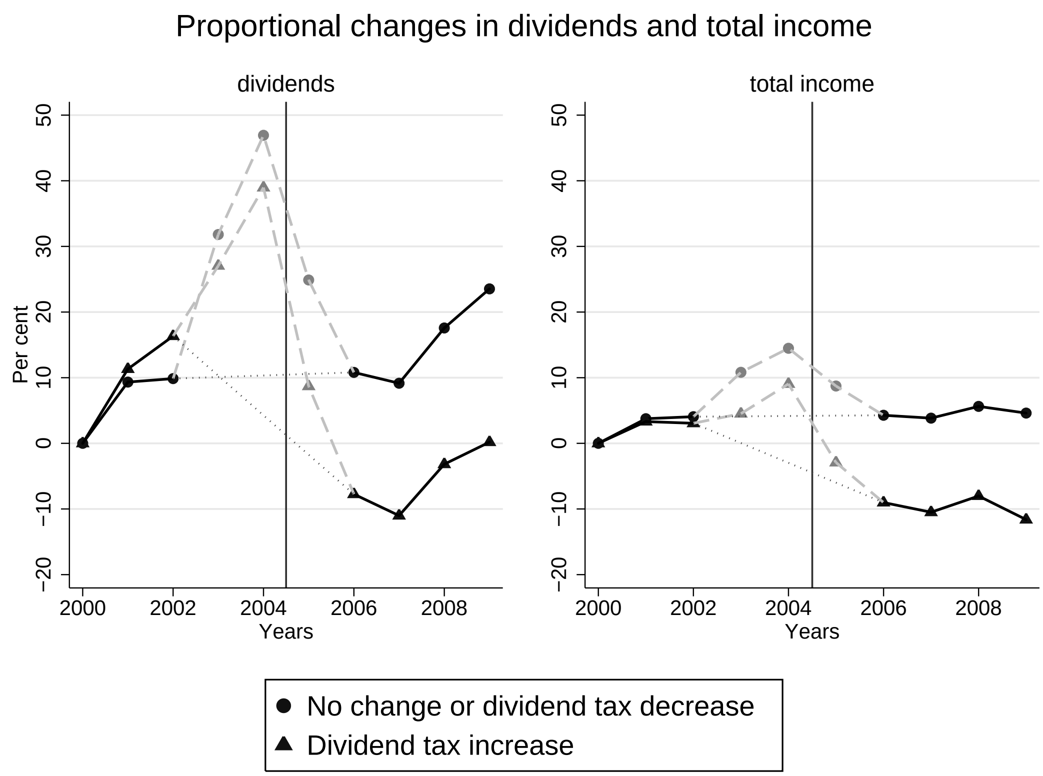


FIGURE 2. Proportional changes in dividends and total income by tax incentives (estimation sample)

The right-hand side of Figure 2 shows that total income increased very similarly in both groups before the anticipation period (2000-2002). This ensures that our estimation results are not biased by differing income trends before the reform. Second, both groups anticipated the future increase in dividend taxation by increasing total income withdrawn from the firm in 2003-2004. Third, compared to owners with no tax changes or a dividend tax decrease, owners with a dividend tax increase decreased their total income after the reform. This implies that the dividend tax reform also affected the total income withdrawn from the firm.

In summary, Figure 2 indicates that owners responded to the dividend tax reform of 2005 according to the tax incentives. Dividends and total income decreased among those who faced a predicted dividend tax increase, compared to owners with no changes or a dividend tax decrease. Changes in dividends could be caused by both real responses and income-shifting. However, changes in total income suggest that at least part of the effect stemmed from real responses. Nevertheless, in order to identify real responses and income-shifting separately, we need to estimate the model with simultaneous changes in both income-shifting and real-term incentives.

3.6. Main Results

3.6.1. ETI and income-shifting. Table 2 presents ETI estimates for wage income and dividend income (gross income subject to taxation) for a single difference between 2002-2007. Columns (1)-(3) show the results for dividends, and columns (4)-(6) present wage income elasticities with the full set of control variables.

For dividends, the standard ETI model in column (1) gives average net-of-tax rate elasticity of over 1.6, which can be considered large. However, as income-shifting possibilities between dividends and wages are particularly relevant for owners of privately held corporations in Finland, we need to add the net-of-tax rate of wages into the model in order to more rigorously analyze tax responsiveness.

Columns (2) and (3) imply that a significant part of the overall behavioral response of dividends is due to income-shifting between the tax bases. Column (2) shows that the cross-elasticity of dividends with respect to the net-of-tax rate on wages is almost -1.5 and statistically significant. However, simply adding the tax rate of another tax base

VARIABLES	(1) lnZ_D	(2) lnZ_D	(3) lnZ_D	(4) lnZ_W	(5) lnZ_W	(6) lnZ_W
$ln(1-t_W)$		-1.468***		0.042	0.316	-0.093
		(0.376)		(0.306)	(0.355)	(0.300)
$ln(1-t_D)$	1.649***	1.989***	0.521*		-0.409***	
	(0.123)	(0.163)	(0.297)		(0.139)	
$[ln(1-t_D)-ln(1-t_W)]$			1.468***			-0.409***
			(0.376)			(0.139)
Income spline	Yes	Yes	Yes	Yes	Yes	Yes
Base-year controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm controls	Yes	Yes	Yes	Yes	Yes	Yes
1st stage F-test for $ln(1-t_W)$		163.69		226.97	172.00	172.00
1st stage F-test for $ln(1-t_D)$	877.02	601.02	601.02		548.66	
1st stage F-test for $[ln(1-t_D)-ln(1-t_W)]$			334.74			333.09
Observations	14,003	14,003	14,003	12,135	12,135	12,135

Notes: Robust standard errors in parentheses. Estimates weighted by total income.*** p<0.01, ** p<0.05, * p<0.1

TABLE 2. ETI estimates for wages and dividends, 2002-2007

to the ETI model for dividends does not change the baseline interpretation of its own net-of-tax rate elasticity, which still reflects both real and income-shifting responses.

Adding the difference of the instrumented net-of-tax rates to the model changes the interpretation of the dividend tax rate parameter. Now the coefficient for dividend net-of-tax rate only includes the real response, which is estimated to be 0.5 and weakly significant. In terms of identifying the income-shifting response, columns (2) and (3) use the same tax rate variation, which gives the same estimates for the income-shifting component. The main difference in the two approaches is the separate estimate for the real response in column (3).

For wages, columns (4)-(6) show that the only statistically significant effect is the income-shifting response, which is estimated to be around -0.4. The wage net-of-tax rate coefficient is insignificant in every specification. The results for the wage income regressions highlight that a tax base might be responsive to tax rates in other tax bases even when its own tax rate elasticity is zero or insignificant.[18]

In summary, the results in Table 2 show that income-shifting can have substantial impact on the behavior of individuals with income-shifting possibilities. This result is in line with previous studies from both the US (e.g. Slemrod 1995, Gordon and Slemrod 2000, Saez 2004) and the Nordic Countries (e.g. Pirttilä and Selin 2011, le Maire and Schjerning 2013). In addition, we find that different tax bases react differently to tax rate changes, both in terms of real and income-shifting responses. We further interpret and discuss our results in Section 3.8.

3.6.2. Real response estimations. As discussed before in Section 3.3, the estimated real response components in Table 2 might not fully reflect the actual real effort or the productivity of the owner. Real responses in the ETI model might be "contaminated" with other tax avoidance measures, such as private consumption within the firm or fringe benefits.

We estimate the net-of-tax rate elasticities for income components that are broader than the separate tax bases for wages and dividends in order to assess the real component in a more diverse manner. First, we analyze the elasticity of total income (wages + dividends). The conclusions based on the estimates in columns (1) and (2) in Table 3 are similar as before. Dividends also appear to affect real behavior, whereas the real elasticity of wages is close to zero and insignificant. The point estimate for dividend tax elasticity is 0.7, which can be considered relatively large.

[18] We also study the heterogeneity of the results across income groups. Following the approach of Gruber and Saez (2002), we divide the sample to four equal-sized quantiles based on base-year wages and dividends. In summary, we find that the real elasticity is smaller and income-shifting elasticity is larger for high-income owners, and vice versa. However, due to the relatively small number of observations in the subsamples, these elasticity estimates are imprecisely measured.

VARIABLES	(1) $ln(Z_D + Z_W)$	(2) $ln(Z_D + Z_W)$	(3) *ln(net profit)*	(4) *ln(net profit)*	(5) *ln(turnover)*	(6) *ln(turnover)*
$ln(1 - t_W)$		0.086		0.206		0.178
		(0.172)		(0.345)		(0.313)
$ln(1 - t_D)$	0.694***		0.335**		0.293*	
	(0.076)		(0.169)		(0.151)	
Income spline	Yes	Yes	Yes	Yes	Yes	Yes
Base-year controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm controls	Yes	Yes	Yes	Yes	Yes	Yes
1st stage F-test for $ln(1 - t_W)$		183.29		407.85		409.66
1st stage F-test for $ln(1 - t_D)$	805.69		850.48		791.86	
Observations	14,010	14,010	13,507	13,507	13,018	13,018

Notes: Robust standard errors in parentheses. Estimates weighted by total income.*** p<0.01, ** p<0.05, * p<0.1

TABLE 3. Real response estimations, 2002-2007

However, there are still various ways in which owners can affect the amount of reported total income subject to taxation. For example, as shown in Figure 2 above, owners anticipated the dividend tax increase in 2005 by increasing both reported dividends and total income before the reform. Therefore, we also estimate the tax elasticities of firm-level income components that are arguably less subject to tax avoidance and intertemporal shifting of reported income.

First, we estimate the elasticity of net profits with respect to both net-of-tax rates. Net profits are defined as turnover plus other income of the firm minus all costs except wages. Thus net profits include, for example, sales, capital gains and irregular earnings. Second, we estimate the net-of-tax rate elasticity of the turnover of the firm. Importantly, compared to total income withdrawn from the firm, both net profits and turnover also include retained earnings and other income not withdrawn from the firm in the current period.

Both of these variables reflect the real effort of the owner. The firms in our estimation sample are relatively small in terms of the number of employees (median no. of employees is 3 in 2007). Thus the owner contributes significantly to the overall output of the firm. Furthermore, the tax elasticities of these types of firm-level income components with respect to owner-level tax rates are rarely analyzed in public finance literature.

Net profits are significantly responsive to dividend taxes, but the point estimate is half the size of the total income elasticity (column (3)). The point estimate for wage tax elasticity increases, but it is still statistically insignificant (column (4)). Columns (5) and (6) present the elasticity estimates for the turnover of the firm. The estimates are similar, but somewhat smaller than before with net profits. This indicates that dividend taxes affect the productivity of the firm and the effort of the owner in a statistically significant manner, whereas wage taxes do not.

However, it is worth noting that the size of the income component might also affect the estimates. As the underlying tax rate variation is the same in all specifications, broader tax bases have smaller elasticities if the absolute behavioral response is the same for different income components. Thus differences in the elasticity estimates for different income types might not be solely driven by differences in the opportunities to avoid taxes.

Overall, the results imply that even though income-shifting between tax bases accounts for a large proportion of the elasticity, the responses along the real margin might still be non-negligible at the same time. Thus for the policy maker, this requires weighting the possible advantages (or disadvantages) stemming from real responses with the costs of avoiding taxes by income-shifting.

Finally, in addition to effort responses, dividend and wage taxes might affect firm-level real investment decisions. Investment responses might cause additional welfare effects, which are not included in our model of the excess burden. However, in order to thoroughly analyze the welfare effects of investments, we would need a well-defined dynamic model. Nevertheless, we address this issue by characterizing investment responses. Similarly as in Table 3, we regress the change in the fixed assets of the firm[19] with the changes in the net-of-tax rates on wages and dividends. In the fixed assets estimations, the coefficient for dividend taxes is positive and statistically significant (0.33 (0.17)). For wage taxes, the estimate is very close to zero and insignificant (0.04 (0.37)). These results are well in line with the turnover and net profits estimations, and highlight that dividend taxes might also significantly affect the real economic decisions of the owners.

3.7. Alternative specifications and robustness checks

3.7.1. Bunching at kink points. Examining taxable income distributions near the kink points of the piecewise linear income tax schedule provides a visual and robust method to analyze the overall ETI. The bunching method provides a local alternative to the first-differences approach, and allows us to estimate behavioral responses using cross-sectional variation in tax rates. This avoids some of the critical issues in first-differences estimation and net-of-tax rate instruments, such as potential mean reversion. Intuitively, similar general conclusions from the bunching analysis would support our main results based on panel data regressions.

A seminal contribution by Saez (2010) shows that under normal preferences, we should observe an excess mass of individuals clustering at the point in the income distribution where the marginal tax rate exhibits a discontinuous jump if significant behavioral responses occur. More formally, Saez (2010) shows that the local ETI is proportional to the excess density mass around the kink point:

[19] Fixed assets include, for example, machinery and equipment, plants and buildings, and research and development expenses and other long-term expenses.

$$e \backsimeq \frac{b(k)}{k * ln((1-\tau_1)/(1-\tau_2))} \tag{3.7.1}$$

In equation (3.7.1), $b(k)$ is the excess mass at the kink point k, and $(1-\tau_1)$ and $(1-\tau_2)$ denote the net-of-tax rates below and above k, respectively. Empirically, $b(k)$ is estimated by comparing the observed income distribution at the kink point to a counterfactual income distribution, representing the shape of the distribution in the absence of a change in the marginal tax rate at k.

Intuitively, given the size of the change in the net-of-tax rate around the kink point, the implied elasticity is larger the more behavioral responses occur and more bunching is observed at the kink. Also, with given excess bunching, the elasticity is smaller the bigger is the difference between the tax rates on both sides of the kink. A more detailed theoretical background for the bunching approach and the estimation procedure are presented in the end of the Appendix.

Figure 3 shows the distributions of dividend income around the kink point of flat-taxed dividends in 2002 and 2007. The figure presents dividend income relative to the kink for each owner within +/- 5,000 € of the kink in bins of 100 €. Dividend income below the kink is taxed at the flat tax rate. Dividends exceeding the kink are taxed progressively. Thus for many owners, the flat-tax kink point induces large changes in the marginal tax rate on dividends. On average, the increase in the MTR on dividends at the kink is 13 percentage points in 2002, and 19 percentage points in 2007.

Figure 3 indicates clear bunching at the flat-tax kink point. A large proportion of the owners are located very close to or exactly at the kink point. This strongly supports the earlier conclusions that owners are responsive to marginal tax rates on dividends, and that the dividend tax base is clearly responsive to its marginal tax rate. We approximate the local ETI of dividend income at the kink point using the average

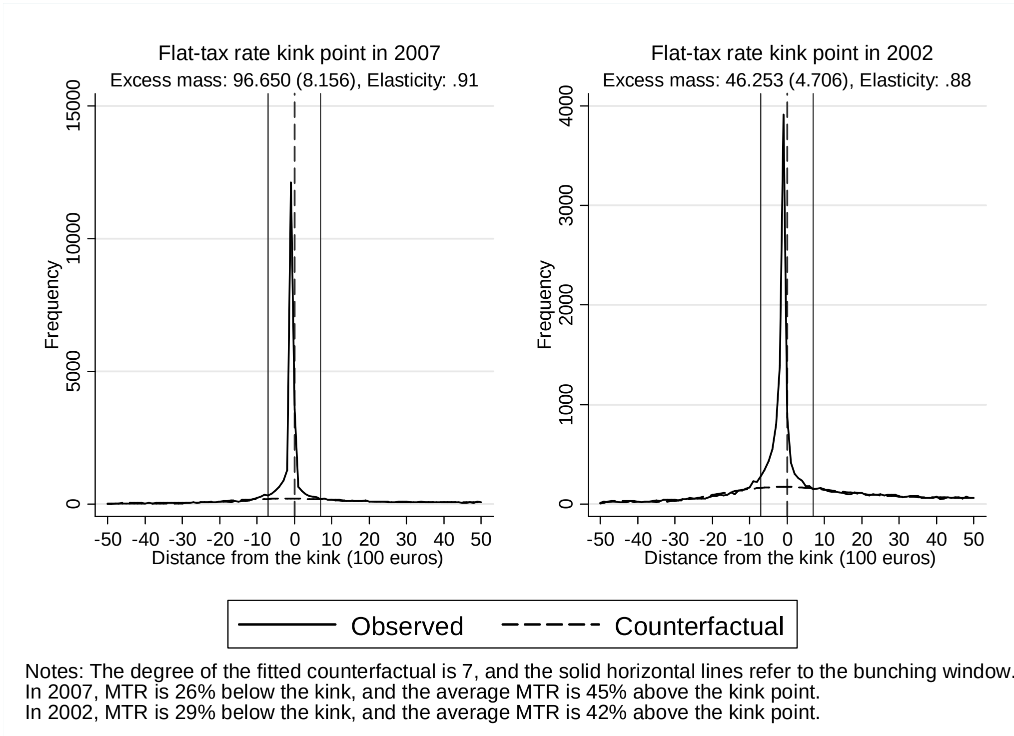


FIGURE 3. Dividend income distribution around the flat-tax rate kink point, years 2007 (left) and 2002 (right)

marginal tax rate above the kink for owners within the bunching window. We estimate the ETI to be around 0.9 and statistically significant both in 2002 and 2007.

There are a few aspects that are worth noting when interpreting Figure 3. First, the flat-tax kink point is not the same for all owners in terms of euros, as the amount corresponding to the 9% imputed return on the net assets of the firm obviously varies among different owners. However, Figure 3 implies that owners are very aware of their individual kink points, as there is no other explicit reason to locate at the kink except the discontinuous change in the tax rate. Second, the size of the change in the marginal tax rate on dividends at the kink point also varies among owners, as the marginal tax rate on dividends exceeding the kink depends on the total sum of progressively taxed income (wages and earned income from other sources).

The bunching method identifies the overall effect of the increase in the MTR on dividends close to the kink, not taking into account potential changes in behavior elsewhere in the income distribution or in other tax bases. Therefore, the bunching approach does not enable us to explicitly identify separate estimates for real elasticity and income-shifting. Nevertheless, the bunching evidence clearly shows that dividend tax rates induce notable behavioral responses among the owners.

In addition, we conduct an indirect bunching analysis for wages. The exact location in the taxable income distribution is what matters in terms of bunching at kink points. Thus it is not relevant to analyze only the distribution of wages from the firm, as other progressively taxed income also affects the location of the owner in the taxable income distribution. For simplicity, in the bunching analysis, we only include owners who do not receive wages or other earned income outside the firm. Nevertheless, the results are similar when we include all the owners in the data set.

Figure 4 presents the distributions of earned income relative to different kink points in the marginal tax rate schedule for 2002 and 2007 (+/- 5,000 € in bins of 100 €). The figure shows that there is no statistically significant excess bunching at the kink points of the earned income tax schedule. For the sake of brevity, Figure 4 presents only 3 kink points in both years, but the result of no significant bunching holds for all kink points.

The evidence from the wage tax rate kink points suggests that owners do not react actively to marginal wage tax rates, which is in line with the low wage elasticity estimates presented before. Compared to the first-differences analysis, the cross-sectional bunching approach is not sensitive to the size of the change in the marginal tax rate between t and $t + k$. As changes in wage tax rates over time have been modest in

2002-2007, this might affect the results in Section 3.6.1. Nevertheless, both of these methods suggest low responsiveness of wage income to the marginal tax rate on wages.

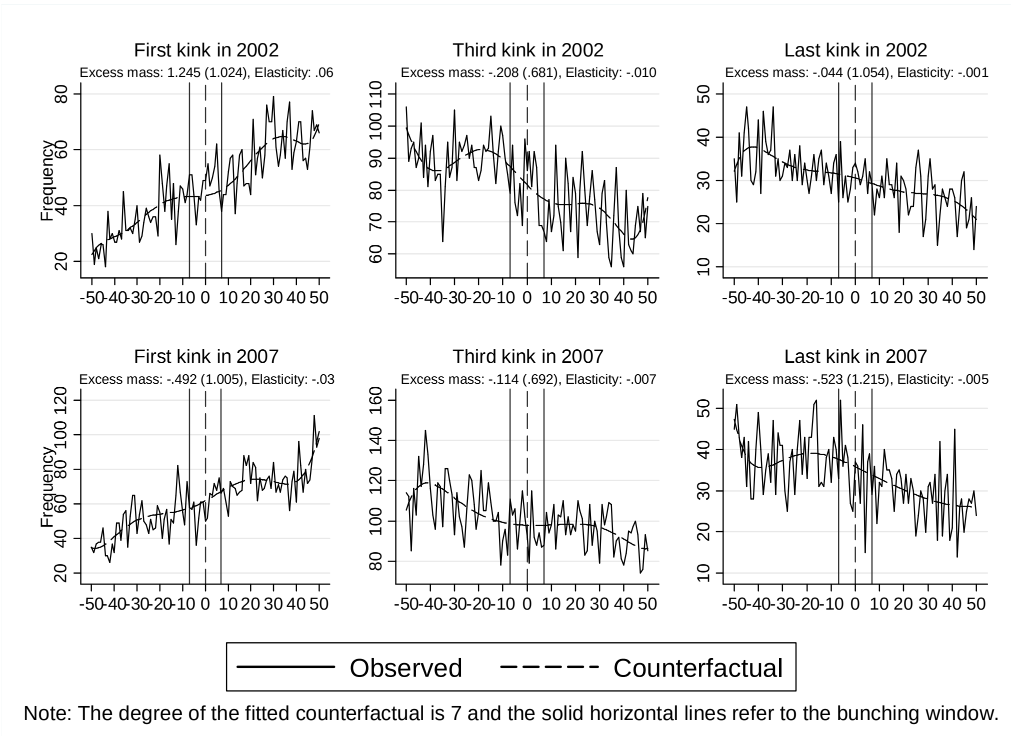


FIGURE 4. Taxable earned income distributions relative to different kink points, years 2002 (above) and 2007 (below)

In summary, the bunching analysis supports the result that dividends are more responsive to tax rates than wages. We find clear bunching at the flat-tax kink point for dividends, whereas the earned income tax rate schedule appears not to induce significant behavioral responses. However, studying excess bunching of dividends and wages does not give explicit information on the extent of income-shifting between the tax bases. Nevertheless, the bunching evidence indicates that owners are very aware of the flat-tax kink point. Given the ample possibilities to shift income between wages and dividends, this tentatively suggests that the dividend tax kink also affects wage income and the overall composition of total income.

3.7.2. Robustness checks and sensitivity analysis. We estimate several different specifications of our estimable equation (3.3.5) in order to assess the robustness and sensitivity of our baseline results. The results for these estimations are presented in Tables A4 and A5 in Appendix. In general, our results and main conclusions are robust to changes in the empirical specification.

In column (1) of Table A4 we estimate the model for dividends, wages and the turnover of the firm without any control variables. The results without controls are approximately similar to those with controls. This tentatively indicates that non-tax-related changes in income do not significantly bias the results. In other words, identification appears not be very sensitive to the selected individual and firm-level controls.[20]

Following Gruber and Saez (2002), our baseline estimates are weighted by total income. Column (2) in Table A4 shows the unweighted estimates, which are very similar to the weighted estimates. Columns (3)-(6) show the results with different variations of income cut-offs. All of these results are statistically equivalent to our baseline model. However, the point estimates vary somewhat depending on the income cut-offs.

In addition, we test how our modifications to the predicted NTR instrument affect our results. First, including the YEL/TEL dummy variable as an additional instrument in the wage regression does not have a significant effect on the point estimates. However, it improves the precision of the estimation. Second, the prediction of firm net assets based on pre-reform observed characteristics affects the main results for dividends,

[20] As an additional robustness check, we add 10-piece splines of firm-level income and asset variables as control variables in order to more rigorously control for the possibility that changes in individual income and firm-level characteristics are connected. This might be a concern because firm net assets, which also reflect the size of the firm, greatly affect changes in the marginal tax rates on dividends. However, adding firm-level splines does not significantly affect the results. Nevertheless, adding additional splines increases precision.

compared to NTR instruments defined with fixed net assets from the base year 2002. Without predicting the net assets, the F-test for the NTR instrument is very low. This indicates that the NTR instrument poorly predicts the changes in tax incentives without taking into account the (counterfactual) growth of net assets, which greatly affects the realized MTR on dividends.

Finally, Table A5 shows the results for longer-run effects, namely for the years 2002-2008 and 2002-2009. The estimates suggest that the income-shifting effect somewhat decreases in the longer run. Otherwise the results are similar to those in our baseline model. This also implies that anticipation of the 2005 tax reform does not significantly affect the results, as the estimates for longer time periods imply qualitatively similar results as our baseline model.

3.8. Discussion

In this paper we show how including income-shifting between tax bases in the widely-used elasticity of taxable income (ETI) framework affects the interpretation of ETI as the overall measure of tax efficiency. Based on previous ETI literature, we build an empirical model that identifies both real responses and income-shifting. As an empirical example, we analyze real responses and income-shifting between wages and dividends among the owners of privately held corporations in Finland. To our knowledge, this paper is the first to explicitly separate these elasticities using a well-defined empirical model and detailed micro-level panel data and tax reforms.

As shown by Feldstein (1995, 1999), ETI summarizes the overall deadweight loss of income taxation. The source of the behavioral response is irrelevant as long as individuals optimize such that the marginal cost of "creating" taxable income through different margins equals the net-of-tax rate of the tax base.

However, tax avoidance through income-shifting might distort this line of thought for at least two reasons. First, if part of the behavioral response is due to income-shifting between tax bases, the shifted income is usually also taxed. Thus not all of the overall response is necessarily a deadweight loss. Second, the real social costs associated with income-shifting might be low, for example due to fiscal externalities related to payments to tax consultants. This further decreases the efficiency loss. In the extreme case in which income-shifting induces no social costs at the margin, the income-shifting response is only a re-allocation of resources between individuals and the government with no welfare losses (Chetty 2009).

Our results show that income-shifting accounts for a large proportion of the overall behavioral response among the owners of privately held businesses in Finland. Over two thirds of the overall dividend income response is due to income-shifting. For wages, the only statistically significant response comes from the income-shifting margin.

What do the results from our empirical example case imply in terms of the excess burden analysis? Applying the welfare loss formulas (3.2.9) and (3.2.10) presented in Section 3.2.3, we can evaluate the marginal excess burden both in the standard ETI framework and the income-shifting model. We approximate the marginal excess burden using the average real and income-shifting elasticities, and the average marginal tax rates on dividend income and wage income (using post-reform values for the whole data set).

Using the standard ETI framework and the point estimate for the average overall dividend elasticity in column (1) of Table 2, we approximate the marginal excess burden of dividend taxation to be around 0.9. When separating the income-shifting effect and using the average estimates in column (3) of Table 2, the marginal excess burden halves to 0.4. Thus the standard ETI analysis for the dividend tax base notably overestimates the deadweight loss, and simply taking into account the fact that the shifted income is also taxed significantly decreases the evaluated efficiency loss.

Furthermore, if we assume that income-shifting is purely transferring resources in the economy with zero social costs, the marginal excess burden of dividend taxes reduces to the welfare loss induced solely by the real behavioral effect. Using the estimate for real responses in the dividend tax base model in column (3) of Table 2, we approximate the marginal deadweight loss to be around 0.3.

However, our estimate for the real response in the single tax base model might be upwards-biased. The real response estimate might include other tax avoidance channels. For example, owners can shift reported personal income intertemporally between different time periods using retained earnings.

We use broader firm-level income components (turnover and net profits) that are less subject to tax avoidance in order to examine the real effects more diversely. In addition to income withdrawn from the firm by the owner, turnover and net profits include retained earnings, and therefore capture the effect of tax rates on effort and productivity in a more comprehensive manner. The results show that both net profits and the turnover of the firm appear to be responsive to dividend taxes, which supports the view that the real responses are non-negligible. In contrast, the effect of wage taxes does not statistically differ from zero.

Overall, we find that the dividend tax base is more responsive to taxes than the wage tax base. Also, firm-level income components are more responsive to dividend taxes compared to wage taxes. There are plausible explanations for these findings. First, the variation in tax rates is larger for dividends, both over time and between income tax brackets. If there are underlying optimization frictions, the owners would respond more to larger changes in tax rates (see e.g. Chetty 2012 and Kleven and Schultz 2013).

Second, there might be practical differences between the two income tax bases. Decisions on dividend distributions are usually made only once or a few times within a year. In contrast, wages are normally paid on a monthly or weekly basis. The infrequent nature of the decision-making process might make dividend income more responsive to taxes. Also, owners might be more aware of the dividend tax rate and the dividend tax rate kink points in the Finnish system, as long as they are aware of the net assets of their firm. In contrast, the effective marginal wage tax rate schedule including many deductions and tax credits might be less transparent.

Third, the return on invested income could be inherently more elastic than the compensation for working. This implies that dividends are simply more responsive to tax rate changes, given the size of the change in incentives and the transparency of the tax system.

In summary, this paper highlights the importance of separating different ways of responding to changes in income tax rates. In terms of welfare analysis, we show that the distinction between income-shifting and real elasticity components can have substantial effect on policy conclusions. This approach is applicable to many other reforms in other countries. For example, the Tax Reform Act of 1986 in the US drastically decreased marginal personal tax rates of high-income earners, and induced notable incentives to shift income from the corporate tax base to the personal tax base. With separate real and income-shifting elasticities, we could outline the true welfare effect of reducing personal income tax rates in the US.

In the empirical part, we show that income-shifting and real elasticities can be estimated separately by including the difference of the relevant tax rates into the model. In addition to clear income-shifting responses, our analysis highlights that it is important to estimate elasticities separately for all relevant tax bases and tax rates. In theory, we have no explicit reason to assume symmetric responses between different tax bases or tax rates. In fact, we find that in Finland the dividend tax base is notably more elastic than the wage tax base, both at the real and income-shifting margin.

Furthermore, the government cannot easily affect the real response margin of individuals, as real responses reflect deeper behavioral parameters, such as the opportunity cost of working (Piketty et al. 2014, Slemrod 1995). However, opportunities for tax avoidance are more under the control of the policy maker. For example, limiting the legal possibilities to avoid taxes presumably decreases income-shifting activity. In the extreme case, setting $\tau_D = \tau_W$ removes all income-shifting incentives in our example, and the government can set tax rates based on real responses alone (assuming no other forms of tax avoidance or evasion).

However, there are also reasons not to set equal tax rates. Within a simple Ramsey framework, it would be optimal to tax less the tax base with the larger elasticity (Piketty et al. 2014). Based on our results, this would imply that the optimal policy is to set $\tau_D < \tau_W$, at least when income-shifting possibilities are absent or restricted.

Nevertheless, it might be that different tax rates and the possibility to shift income between tax bases increase overall entrepreneurial activity and effort in the long run. Theoretically, this refers to a model in which the cost functions of real effort and income-shifting are not separable, and thus separate elasticity parameters for income-shifting and real behavior cannot be empirically identified in our framework. In this generic case, the policy maker needs to balance between the inefficiency and revenue losses induced by income-shifting, and the potential long-run efficiency gains induced by setting differential tax rates and allowing for income-shifting.

Bibliography

[1] Alstadsæter, A. and M. Jacob. (2012). Dividend taxes and income shifting. FAccT Center Working Paper No. 07/2012.

[2] Bastani, S. and H. Selin. (2014). Bunching and non-bunching at kink points of the Swedish tax schedule. Journal of Public Economics, 109: 36–49.

[3] Blomquist, S. and H. Selin. (2010). Hourly wage rate and taxable labor income responsiveness to changes in marginal tax rates. Journal of Public Economics, 94(11-12): 878–889.

[4] Chetty, R. (2012). Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply. Econometrica, 80(3): 969–1018.

[5] Chetty, R. (2009). Is the taxable income elasticity sufficient to calculate the deadweight loss? The implications of evasion and avoidance. American Economic Journal: Economic Policy, 1(2): 31–52.

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

[7] Devereux, M., L. Liu and S. Loretz. (2014). The elasticity of corporate taxable income: New evidence from UK tax records. American Economic Journal: Economic Policy, 6(2): 19–53.

[8] Feldstein, M. (1999). Tax avoidance and the deadweight loss of the income tax. Review of Economics and Statistics, 81(4): 674–680.

[9] Feldstein, M. (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.

[10] Goolsbee, A. (2000). What happens when you tax the rich? Evidence from executive compensation. Journal of Political Economy, 108: 352–378.

[11] Gordon, R. and J. Slemrod. (2000). Are 'real' responses to taxes simply income shifting between corporate and personal tax bases?, in J. Slemrod (ed.), Does Atlas Shrug? The Economic Consequences of Taxing the Rich. New York: Russell Sage Foundation and Harvard University Press, 2000: 240–280.

[12] Gruber, J. and E. Saez. (2002). The elasticity of taxable income: evidence and implications. Journal of Public Economics, 84: 1–32.

[13] Harju, J. and T. Kosonen. (2013). The impact of tax incentives on the economic activity of entrepreneurs. CESifo Working Paper Series 4259.

[14] Harju, J. and T. Matikka. (2012). Entrepreneurs and income-shifting: Empirical evidence from a Finnish tax reform. Government Institute for Economic Research Working Papers 43.

[15] Kari, S., H. Karikallio and J. Pirttilä. (2008). Anticipating tax changes: Evidence from the Finnish corporate income tax reform of 2005. Fiscal Studies, 29(2): 167–196.

[16] Kleven, H. and E. Schultz. (2013). Estimating taxable income responses using Danish tax reforms. Forthcoming, American Economic Journal: Economic Policy.

[17] Lindsay, L. (1987). Individual taxpayer response to tax cuts: 1982-1984, with implications for the revenue maximizing tax rate. Journal of Public Economics, 33(2): 173–206.

[18] le Maire, D. and B. Schjerning. (2013). Tax bunching, income shifting and self-employment. Journal of Public Economics, 107: 1–18.

[19] Piketty, T., E. Saez and S. Stantcheva. (2014). Optimal taxation of top labor incomes: A tale of three elasticities. American Economic Journal: Economic Policy, 6(1): 230–271.

[20] Pirttilä, J. and H. Selin. (2011). Income shifting within a dual income tax system: Evidence from the Finnish tax reform of 1993. Scandinavian Journal of Economics, 113(1): 120–144.

[21] Romanov, D. (2006). The corporation as a tax shelter: Evidence from recent Israeli tax changes. Journal of Public Economics, 90(10-11): 1939–1954.

[22] Saez, E. (2010). Do taxpayers bunch at kink points? American Economic Journal: Economic Policy, 2(3): 180–212.

[23] Saez, E. (2004). Reported income and marginal tax rates, 1960-2000: Evidence and policy implications. NBER Working Paper 10273.

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

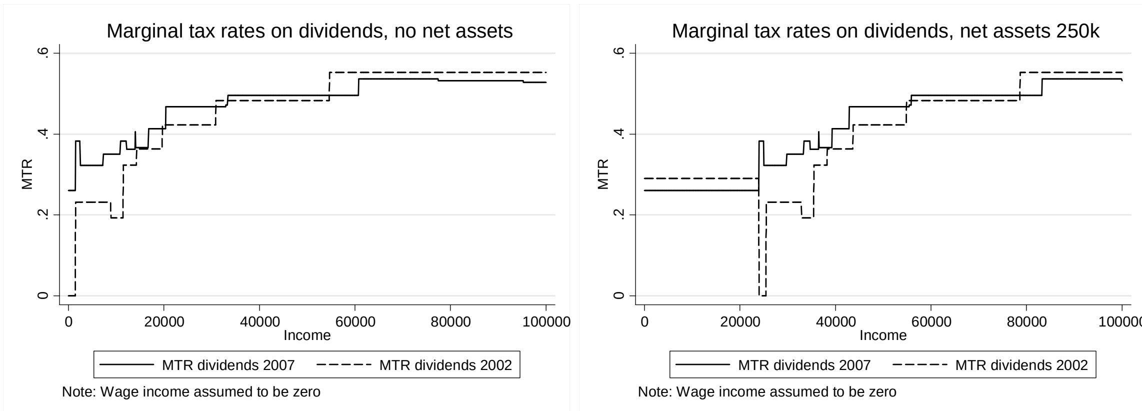
[25] Sivadasan, J. and J. Slemrod. (2008). Tax law changes, income-shifting and measured wage inequality: Evidence from India. Journal of Public Economics, 92(10-11): 2199–2224.

[26] Slemrod, J. (1995). Income creation or income shifting? Behavioral responses to the Tax Reform Act of 1986. American Economic Review Papers and Proceedings, 85(2): 175–180.

[27] Slemrod, J. and C. Gillitzer. (2014). Tax systems. Cambridge: The MIT Press, 2014.

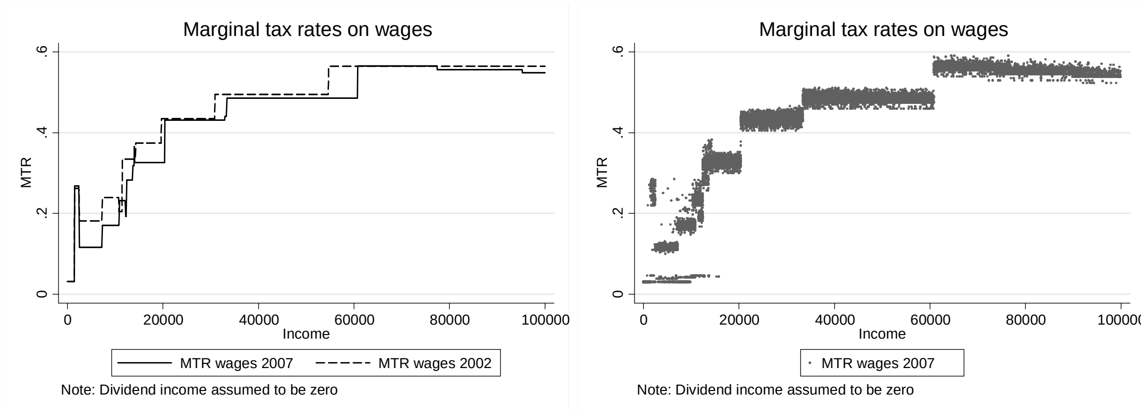
[28] Weber, C. (2014). Toward obtaining a consistent estimate of the elasticity of taxable income using difference-in-differences. Forthcoming, Journal of Public Economics.

Appendix



Notes: MTR on dividends includes corporate taxes paid on withdrawn dividends (after 2005) and all automatic deductions and allowances on dividend income. Progressively taxed dividends include central government taxes and average municipal taxes.

Figure A1. Average marginal tax rates on dividends in 2002 and 2007. No net assets (left-hand side), net assets of 250,000 € (right-hand side)



Notes: Left-hand side: Average MTR includes central government taxes, average municipal taxes and all automatic tax deductions and exemptions. MTR also includes social security contributions levied on wage income as well as firm-level social security contributions.

Right-hand side: MTR includes central government taxes, individual municipal taxes and individual tax deductions and exemptions. MTR also includes social security contributions levied on wage income as well as firm-level social security contributions.

Figure A2. Average marginal tax rates on wages in 2002 and 2007 (left-hand side). Marginal tax rates on wages in 2007, including individual variation in the municipal tax rate (right-hand side)

	MTR on wages		MTR on dividends (no net assets)		MTR on dividends (net assets 250k)		MTR on dividends (net assets 1000k)		MTR on dividends (net assets 5000k)	
Income	*2002*	*2007*	*2002*	*2007*	*2002*	*2007*	*2002*	*2007*	*2002*	*2007*
5,000	18.1	11.6	23.1	32.3	29.0	26.0	29.0	26.0	29.0	26.0
10,000	23.9	17.0	19.3	35.1	29.0	26.0	29.0	26.0	29.0	26.0
15,000	37.4	32.6	36.3	36.6	29.0	26.0	29.0	26.0	29.0	26.0
20,000	43.4	32.6	42.3	41.3	29.0	26.0	29.0	26.0	29.0	26.0
25,000	43.4	43.1	42.3	46.7	0	32.3	29.0	26.0	29.0	26.0
30,000	43.4	43.1	42.3	46.7	23.1	35.1	29.0	26.0	29.0	26.0
35,000	49.4	48.5	48.3	49.5	19.3	36.2	29.0	26.0	29.0	26.0
40,000	49.4	48.5	48.3	49.5	36.3	41.3	29.0	26.0	29.0	26.0
45,000	49.4	48.5	48.3	49.5	42.3	46.7	29.0	26.0	29.0	26.0
50,000	49.4	48.5	48.3	49.5	42.3	46.7	29.0	26.0	29.0	26.0
55,000	56.4	48.5	55.3	49.5	48.3	46.7	29.0	26.0	29.0	26.0
60,000	56.4	48.5	55.3	49.5	48.3	49.5	29.0	26.0	29.0	26.0
65,000	56.4	56.5	55.3	53.7	.48.3	49.5	29.0	26.0	29.0	26.0
70,000	56.4	56.5	55.3	53.7	48.3	49.5	29.0	26.0	29.0	26.0
75,000	56.4	56.5	55.3	53.7	48.3	49.5	29.0	26.0	29.0	26.0
80,000	56.4	55.6	55.3	53.2	55.3	49.5	29.0	26.0	29.0	26.0
85,000	56.4	55.6	55.3	53.2	55.3	53.7	29.0	26.0	29.0	26.0
90,000	56.4	55.6	55.3	53.2	55.3	53.7	29.0	32.3	29.0	40.5
95,000	56.4	55.6	55.3	53.2	55.3	53.7	29.0	35.1	29.0	40.5
100,000	56.4	54.8	55.3	52.8	55.3	53.2	23.1	36.2	29.0	40.5

Notes:

MTR on wages is calculated with dividend income equal to zero, and vice versa. MTR on wages include average municipal taxes, central government income taxes, automatic tax deductions and tax credits and average firm-level social security contributions (3%). MTR on wages does not include pension and health insurance contributions or any deductions based on insurance contributions. MTR on dividends includes corporate taxes on withdrawn dividends (after 2005). MTR on dividends include all automatic tax deductions and tax credits. MTR on progressively taxed dividends include average municipal taxes and central government income taxes. Marginal tax rates are calculated using Stata and the JUTTA microsimulation model.

TABLE A1. Marginal tax rates (MTR) on wages and dividends with different levels of firm net assets, 2002 and 2007 (in nominal euros)

	2002				2007			
Variable	Mean	Median	SD	N	Mean	Median	SD	N
Wages	25,862	21,306	34,688	39,101	30,780	25,615	40,964	52,028
Dividends	25,696	8,750	101,722	39,104	22,015	7,523	83,456	52,045
Total income	51,560	35,242	110,046	39,101	52,800	38,458	95,633	52,028
MTR dividends	0.38	0.37	0.11	39,104	0.36	0.26	0.11	52,045
MTR wages	0.47	0.51	0.11	39,104	0.42	0.47	0.13	52,045
YEL	0.35	0	0.48	39,104	0.54	1	0.50	52,045
Ownership share	0.80	0.70	0.35	39,104	0.73	0.80	0.27	52,045
Male	0.82	1	0.38	39,104	0.82	1	0.38	52,045
Age	48.47	49	10.46	39,104	50.42	51	10.78	52,045
Turnover	1,022,725	232,099	5,847,782	39,104	1,064,023	224,399	8,153,712	52,045
Total assets	697,755	167,336	4,410,689	39,104	855,857	196,591	6,140,952	52,045
Net assets	431,001	93,075	3,836,671	39,104	524,072	108,413	4,034,409	52,045
No. of employees	10.74	3	47.76	39,104	9.74	3	51.52	52,045

Table A2. Descriptive statistics, data (in 2002 euros)

	2002				2007			
Variable	Mean	Median	SD	N	Mean	Median	SD	N
Wages	27,300	25,000	21,207	14,012	28,992	26,546	24,237	14,010
Dividends	21,028	11,301	32,882	14,012	22,251	11,878	33,858	14,010
Total income	48,328	40,738	38,152	14,012	51,243	44,050	41,118	14,010
MTR dividends	0.40	.42	0.10	14,012	0.37	0.26	0.11	14,010
MTR wages	0.48	0.51	0.09	14,012	0.43	0.47	0.12	14,010
YEL	0.62	1	0.49	14,012	0.62	1	0.49	14,010
Ownership share	0.77	0.80	1.02	14,012	0.76	0.85	0.26	14,010
Male	0.84	1	0.37	14,012	0.84	1	0.37	14,010
Age	47.4	48	9.27	14,012	52.4	53	9.27	14,010
Turnover	764,175	265,622	2,652,496	14,012	852,451	267,531	2,732,651	14,010
Total assets	453,071	190,734	1,686,930	14,012	650,201	250,470	2,612,920	14,010
Net assets	268,107	113,133	837,228	14,012	399,598	154,933	1,634,324	14,010
No. of employees	8.94	4	21.31	14,012	8.84	3	23.21	14,010

Table A3. Descriptive statistics, baseline estimation sample (in 2002 euros)

	(1)	(2)	(3)	(4)	(5)	(6)
	No controls	No weights	Small inc limit	Large inc limit	Small △inc limit	Large △inc limit
VARIABLES	lnZ_D	lnZ_D	lnZ_D	lnZ_D	lnZ_D	lnZ_D
$ln(1-t_D)$	0.595*	0.542*	0.860***	0.436	0.753**	0.581**
	(0.317)	(0.290)	(0.259)	(0.331)	(0.307)	(0.293)
$[ln(1-t_D) - ln(1-t_W)]$	1.722***	1.481***	1.359***	1.364***	1.059***	1.535***
	(0.399)	(0.369)	(0.333)	(0.418)	(0.397)	(0.369)
Observations	14,003	14,003	16,935	9,888	10,988	14,879
VARIABLES	lnZ_W	lnZ_W	lnZ_W	lnZ_W	lnZ_W	lnZ_W
$ln(1-t_W)$	-0.274	-0.123	-0.193	-0.158	-0.315	0.011
	(0.308)	(0.299)	(0.273)	(0.347)	(0.296)	(0.298)
$[ln(1-t_D) - ln(1-t_W)]$	-0.691***	-0.420***	-0.423***	-0.391***	-0.594***	-0.328**
	(0.162)	(0.141)	(0.136)	(0.149)	(0.143)	(0.138)
Observations	12,135	12,135	14,342	8,535	9,611	12,870
VARIABLES	*ln(turnover)*	*ln(turnover)*	*ln(turnover)*	*ln(turnover)*	*ln(turnover)*	*ln(turnover)*
$ln(1-t_D)$	0.357**	0.291*	0.297**	0.411**	0.217	0.281*
	(0.160)	(0.149)	(0.143)	(0.171)	(0.169)	(0.146)
Observations	13,018	13,018	15,720	9,134	10,252	13,817
VARIABLES	*ln(turnover)*	*ln(turnover)*	*ln(turnover)*	*ln(turnover)*	*ln(turnover)*	*ln(turnover)*
$ln(1-t_w)$	0.148	0.156	0.266	0.136	0.460	0.042
	(0.329)	(0.306)	(0.272)	(0.360)	(0.337)	(0.305)
Observations	13,018	13,018	15,720	9,134	10,252	13,817

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1 Small income limit (column(3)) = 10,000 € of total income in base-year. Large income limit (4) = 40,000 € of total income in base-year. Small limit in change of income (5) = 25,000 €. Large limit in change of income (6) = 75,000 €. Columns (2)-(6) include all individual and firm-level controls.

Table A4. Robustness checks: Different specifications, 2002-2007

Years 2002-2008						
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	lnZ_D	lnZ_W	$lnTI$	$lnTI$	*ln(turnover)*	*ln(turnover)*
$ln(1-t_w)$		-0.043		-0.015	0.110	
		(0.261)		(0.157)	(0.294)	
$ln(1-t_D)$	0.625**		0.687***			0.207
	(0.250)		(0.083)			(0.161)
$[ln(1-t_D)-ln(1-t_W)]$	1.408***	-0.359**				
	(0.315)	(0.149)				
Observations	12,859	11,012	12,867	12,867	11,840	11,840
Years 2002-2009						
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	lnZ_D	lnZ_W	$lnTI$	$lnTI$	*ln(turnover)*	*ln(turnover)*
$ln(1-t_w)$		0.283		-0.128	0.228	
		(0.292)		(0.163)	(0.309)	
$ln(1-t_D)$	0.396*		0.487***			0.329*
	(0.225)		(0.081)			(0.170)
$[ln(1-t_D)-ln(1-t_W)]$	0.831***	-0.402**				
	(0.292)	(0.156)				
Observations	11,843	9,933	11,851	11,851	10,712	10,712

Notes: Robust standard errors in parentheses. Estimates weighted by total income.

*** p<0.01, ** p<0.05, * p<0.1

Columns (1)-(6) include all individual and firm-level controls

TABLE A5. Robustness checks: Years 2002-2008 and 2002-2009

Bunching at kink points

The intuition behind the bunching analysis is the following: Consider a small increase in the marginal tax rate, $d\tau$, at a point $z = k$. Below the kink point k taxable income z is taxed with a tax rate τ_1, and above the kink point the tax rate is τ_2, such that $\tau_1 < \tau_2$. Assuming individuals with standard preferences as before in Section 3.2, we can denote the fraction of individuals bunching as $B(dz) = \int_k^{k+dz} h_0(z)dz$, where $h_0(z)$ is the pre-reform smooth density function of taxable income. Individuals located within the income interval $(k, k + dz)$ before the tax rate change bunch at k due to the introduction of the kink point. Individuals further up in the income distribution $z > k + dz$ or below k do not move to the kink point. The bunching approach implicitly assumes that individuals in the neighborhood of k are otherwise similar except that they face a different slope in the budget set.

Empirically, the excess mass at the kink point $b(k) = B(dz)/h_0(k)$ is estimated by comparing the actual density around the kink point to a smooth counterfactual density. The counterfactual density describes how the income distribution at the kink would have looked like without a change in the tax rate. Due to imperfect control and uncertainty about the exact amount of income in each year, the usual approach is to use a "bunching window" around k to estimate the excess mass (see Saez 2010 and Chetty et al. 2011). In other words, we compare the density of taxpayers within an income interval $(k - \delta, k + \delta)$ to an estimated counterfactual density within the same income range.

We use the approach of Chetty et al. (2011) and estimate the counterfactual density non-parametrically. We fit a flexible polynomial function to the observed density function, excluding the region around the kink point $[k - \delta, k + \delta]$ from the regression.

First, we group individuals into small income bins, and estimate a bunching regression of the following form

$$c_j = \sum_{i=0}^{p} \beta_i (z_j)^i + \sum_{i=k-\delta}^{k+\delta} \eta_i \cdot \mathbf{1}(z_j = i) + \varepsilon_j$$

where c_j is the count of individuals in bin j, and z_j denotes the income level in bin j. The order of polynomial is denoted by p.

The counterfactual density function is estimated by omitting the bunching window from the regression, $\hat{c}_j = \sum_{i=0}^{p} \beta_i (z_j)^i$. Thus we can express bunching around k as $\hat{B} = \sum_{i=k-\delta}^{k+\delta} c_j - \hat{c}_j$. Finally, the excess mass is calculated as

$$\hat{b}(k) = \frac{\hat{B}}{\sum_{i=k-\delta}^{k+\delta} \hat{c}_j/(2\delta + 1)}$$

As in the earlier literature, parameters δ and p are determined visually and based on the fit of the model. We use a seventh-order polynomial and a bunching window of +/-700 € from the kink point in our baseline estimations. Our conclusions are not very sensitive to the bunching window δ or the degree of polynomial p. As in Chetty et al. (2011), standard errors for $\hat{b}(k)$ are calculated using a bootstrap procedure where we generate a large number of income distributions by randomly resampling the residuals from the bunching regression. The standard errors are defined as the standard deviation in the distribution of $\hat{b}(k)$.

CHAPTER 4

Taxable Income Elasticity and the Anatomy of Behavioral Response: Evidence from Finland[1]

Abstract. This paper uses extensive Finnish panel data to analyze the elasticity of taxable income (ETI). I use changes in flat municipal income tax rates as an instrument for the overall changes in marginal tax rates. This instrument is not a function of individual income, which is the basis for an exogenous instrument in the taxable income model. In general, instruments used in previous studies do not have this feature. Furthermore, I estimate behavioral responses using smaller subcomponents of taxable income, such as working hours, fringe benefits and tax deductions. This "anatomy" of the overall ETI has been rarely studied in the literature. My preferred estimate for the average ETI in Finland is 0.27. The preferred specification includes extensive regional and individual controlling. Subcomponent analysis suggests that neither work effort nor labor supply respond actively to tax changes. In contrast, it seems that fringe benefits and tax deductions might have a more considerable effect.

Keywords: income taxation, elasticity of taxable income, excess burden, labor supply

JEL codes: H21; H24; H31; J22

4.1. Introduction

The elasticity of taxable income (ETI) with respect to the net-of-tax rate (one minus the marginal tax rate) is a key tax policy parameter and an important element in the efficiency analysis of income taxation. The practical significance of ETI is straightforward: it measures how a one percent change in the net-of-tax rate affects taxable income. Intuitively, the more elastic taxable income is, the larger the behavioral response to a tax reform will be, in terms of a change in the tax base.

Under general conditions, ETI measures the marginal excess burden of income taxation (Feldstein 1995, 1999). In addition to labor supply responses, ETI covers changes

[1]An earlier version of this paper is published in the Government Institute for Economic Research Working Papers series, 55, February 2014.

in, for example, effort and productivity, deduction behavior, tax evasion and tax avoidance. All of these margins are (more or less) important when considering the overall efficiency of a tax system. From this point of view, a large ETI makes a tax increase relatively costly and a tax decrease less costly, and vice versa. Altogether, good knowledge of country-specific ETI is essential when deciding on national tax reforms.

This paper studies ETI using Finnish register data. I use changes in flat municipal income tax rates as instruments for the changes in overall net-of-tax rates. This instrument is not based on individual income, which provides the basis for an exogenous instrument. In addition to average ETI, I study the structure of the overall response by characterizing how the subcomponents of ETI, such as labor supply and deduction behavior, react to changes in tax rates.

Earlier empirical literature has mainly focused on estimating the overall elasticity of taxable income. It is still largely unknown which of the behavioral margins are the most responsive components of the total elasticity. However, detailed knowledge of "the anatomy of behavioral response" (Slemrod 1996) is also useful when designing an income tax system and the detailed structure of tax reforms, especially in the light of minimizing the excess burden of income taxation.[2]

Furthermore, analysis of different subcomponents provides information on the economic nature of the response. It is rather difficult for the policymaker to influence deep individual utility arguments, such as the opportunity cost of working. However, for example, it is easier to influence tax deduction behavior through minor adjustments to regulations.

[2]In previous studies, Blomquist and Selin (2010) estimate the elasticity of the hourly wage rate. Using a Swedish data set, they find a significant wage rate response. Also, Kleven and Schultz (2013) report that capital income components of taxable income are more responsive than earned income in Denmark. Previous literature concerning tax reforms in the United States shows that a large proportion of the behavioral response of high-income individuals has been in the form of tax avoidance via income-shifting rather than real economic behavior (see for example Slemrod 1995, 1996, Gordon and Slemrod 2000, Goolsbee 2000, Saez 2004, and Saez, Slemrod and Giertz 2012).

The source of individual variation in tax rates and the endogeneity of the net-of-tax rate variable are the main issues to focus on when estimating ETI using panel data. Identification requires variation in income tax rates that is different for individuals with otherwise similar income trends. Also, due to the progressive income tax rate schedule, a valid instrument for the net-of-tax rate is usually necessary in order to derive a consistent elasticity estimator. In this study I use variation in municipal-level flat income tax rates for both purposes.

Finnish municipal taxation has appealing features from the point of view of empirical ETI analysis. First, the municipal income tax rate is proportional, which means that it is independent of individual income level. This is the basis for using changes in municipal tax rates as an instrument for changes in progressive individual marginal tax rates. Recent literature highlights that frequently used predicted net-of-tax rate instruments are not necessarily consistent (see for example Blomquist and Selin (2010) and Weber (2014)). These instruments are functions of individual income in the base period, and thus potentially endogenous in a model where changes in taxable income are regressed with changes in the instrumented net-of-tax rate.

Second, different municipalities have changed their tax rates differently in different years. In other words, net-of-tax rates have changed differently for otherwise similar individuals who differ only in location. Moreover, as the municipal income tax rate does not depend on individual income, changes in municipal taxation have an effect on net-of-tax rates throughout the income distribution. This enables me to identify the average elasticity parameter while avoiding some of the usual difficulties in ETI estimation, namely non-tax-related changes in the shape of the income distribution and the mean reversion of income. These issues are particularly troublesome if tax rate variation is concentrated in a single part of the income distribution, such as in the case of tax reforms affecting only high-income earners. Many earlier studies base their estimation strategy on tax rate variation that occurs only at certain income levels.

However, changes in municipal income tax rates are not randomly assigned. Municipalities might change their tax rates based on, for example, previous trends in average taxable income in their jurisdiction. This might affect the validity of the instrument. As a potential solution, the data include a variety of municipal characteristics that I use to control for municipal-level economic circumstances. In addition, I apply different combinations of year and regional fixed effects in the estimable equation, and study the effect of previous income trends on future tax changes in order to assess the exogeneity of the instrument.

To summarize, this study contributes to the ETI literature in three ways: First, I use a novel net-of-tax rate instrument based on changes in flat municipal tax rates, which is uncorrelated with changes in individual income. Second, the differential tax rate variation used in this study covers the entire income distribution. This improves the identification of the average ETI, which is the parameter of main interest in this study. Third, I characterize how different components of taxable income react to tax rate changes. This subcomponent analysis provides information on what the most important behavioral margins are. Studying the structure of the elasticity also describes how much of the response is driven by changes in real-term behavior (hours of work and work effort), and how much is accounted for by other margins (tax deductions, fringe benefits). Different components of ETI are rarely analyzed in the empirical literature.

I estimate the average intensive margin ETI in Finland to be 0.27-0.60, depending on the degree of individual and regional controlling. My preferred estimate is 0.27. The preferred specification includes extensive regional and individual controlling.

In addition, as in many earlier studies, the average ETI is larger for women than for men, and larger for high and low-income individuals than for middle-income earners. Furthermore, I provide clear graphical reduced-form evidence that individuals respond to changes in municipal income tax rates, and show that the potential policy endogeneity of municipal tax rate changes is not driving the results.

Analysis of the subcomponents of taxable income gives tentative evidence that both work effort and labor supply are not very responsive to tax rate changes. However, more irregular components, such as fringe benefits and tax deductions, seem to be more responsive. These imply that a large proportion of the overall ETI is not stemming from changes in labor supply behavior.

The empirical ETI literature has grown substantially following the pioneering studies by Feldstein (1995, 1999). Feldstein (1995) uses panel data to analyze behavioral responses to the 1986 tax reform in the US. He estimates ETI to be large, ranging from 1-3, depending on the specification used. Many studies following Feldstein (1995) focus on improving the elasticity estimation by paying more attention to net-of-tax rate instruments and non-tax-related changes in the income distribution. Along with these modifications, the elasticity estimates decreased markedly compared to those in Feldstein (1995). A wide range of studies report elasticity estimates ranging from 0 to 0.6. For example, the widely cited Gruber and Saez (2002) study reports the ETI of 0.2 for middle-income earners, and 0.6 for high-income earners in the US. An extensive review of earlier empirical results from the US is presented in a recent survey by Saez, Slemrod and Giertz (2012).

More recent papers further study the reliability and consistency of the estimation by utilizing different tax reforms and different net-of-tax rate instruments. This literature underlines that different tax reforms and more consistent estimation strategies yield different estimates than the seminal contribution of Gruber and Saez (2002). In particular, it has been shown that the widely-used predicted net-of-tax rate instruments built on base-year income are not consistent due to potential endogeneity problems. For example, Weber (2014) shows that when using a more robust instrumentation strategy, the ETI estimates are twice as large compared to Gruber and Saez (2002).

A majority of earlier empirical studies estimate ETI using US data sets, while studies concerning European countries and other regions are less common. In particular, there are no earlier Finnish ETI studies available to this day.[3] For other Nordic countries, Blomquist and Selin (2010) estimate ETI of around 0.20 for males and 1 for females in Sweden. In addition, they document positive elasticity estimates for the hourly wage rate, and find statistically significant income effects. For Denmark, Kleven and Schultz (2013) use extensive panel data and many tax reforms to analyze ETI. In general, they obtain modest elasticity estimates, the upper bound of ETI being 0.3. Also, Chetty et al. (2011) report small elasticity estimates using Danish data. For Norway, Aarbu and Thoresen (2001) find only small responses to tax changes. They report that ETI is not significantly different from zero. In a more recent paper, Thoresen and Vattø (2013) report elasticities below 0.1 for Norway.

The rest of the paper is organized as follows: Section 4.2 presents the conceptual framework, including the theoretical background and the empirical model. Section 4.3 describes the Finnish income tax system and recent changes in income taxation. Section 4.4 introduces the data and discusses identification issues. Section 4.5 presents the results, and Section 4.6 concludes.

4.2. Conceptual framework

4.2.1. Taxable income model. The basic idea of the (static) taxable income model is that an individual receives positive utility from consumption c and negative utility from creating and reporting taxable income TI (Feldstein 1999).[4] Following Piketty et al. (2014), the quasi-linear utility function $u(c, TI) = c - h(TI)$ is maximized under the budget constraint $c = TI(1-\tau) + R$. The convex cost function of creating

[3] Pirttilä and Uusitalo (2005) calculate approximate elasticity estimates for Finland. Their results suggest that the ETI is around 0.3.

[4] Within this study, taxable income is regarded as taxable earned income. Taxable earned income is defined as the sum of labor income and taxable non-labor income minus deductions (verotettava tulo). The legal distinction between earned income and capital income in the Finnish income tax system is described in the next section.

taxable income is denoted as $h(TI)$, and $(1-\tau)$ denotes the net-of-tax rate on a linear segment of the tax rate schedule. R denotes virtual income.

Maximization of the utility function with respect to the budget constraint gives the supply function of taxable income of the form $TI = TI((1-\tau))$. Next, consider a marginal decrease in $(1-\tau)$ (i.e. a marginal increase in τ). With a quasi-linear utility model with no income effects, the decreased net-of-tax rate decreases the supply of taxable income. Thus we get the following expression

$$\frac{dTI}{TI} = e\frac{d(1-\tau)}{(1-\tau)} \tag{4.2.1}$$

where e is the elasticity of taxable income.

As in most earlier studies, I assume no income effects. Earlier literature shows that income effects are either insignificant or very small (Saez et al. 2012). Thus in the empirical analysis, ETI is measured by regressing changes in taxable income with changes in the net-of-tax rate.

4.2.2. The marginal excess burden of income taxation and the components of taxable income. As shown in Feldstein (1999), all behavioral responses reflect the inefficiency of the income tax system. The marginal excess burden of the income tax can be expressed in terms of the elasticity of taxable income and the relevant income tax rate even when individuals make various decisions in response to income taxation, such as hours of work, work effort, deduction behavior, education choices and so on. This result holds when agents do not make optimization errors, and income taxation or taxable income do not impose any significant externalities.

Following Chetty (2009), consider an individual who makes a vector of decisions $\{x_1, .., x_n\}$ that all affect total taxable income linearly, additively and separately. In this framework, overall taxable income can be presented as the sum of all behavioral choices,

$\Sigma x_i = TI$. As in Chetty (2009), I Assume that each choice x_i has a convex cost function $g_i(x_i)$. Individual maximizes a utility function of the form $u(c, \Sigma x_i) = c - \Sigma g_i(x_i)$ with respect to $c = \Sigma x_i(1-\tau) + R$.

I follow the standard approach in the literature and in Chetty (2009), and assume that the tax revenue collected by the income tax is returned to the individual as a lump-sum transfer. Thus any behavioral responses to tax rates induce welfare losses. The social welfare function used for analyzing the welfare effect is presented as the sum of individual utility (in the curly brackets) and government tax revenue

$$W = \Big\{(1-\tau)\sum x_i + R - \sum g_i(x_i)\Big\} + \tau \sum x_i \tag{4.2.2}$$

Next, consider a small tax increase $d\tau$. As the individual has optimized his/her bundle of x_i, we can write the marginal excess burden of income taxation in the following form[5]

$$DWL = \frac{dW}{d\tau} = \tau \sum_{i=1}^{n} \frac{dx_i}{d\tau} = \tau \frac{dTI}{d\tau} \tag{4.2.3}$$

Most of the earlier studies focus on estimating the overall average ETI. As underlined in Feldstein (1999), the substitution elasticities for different choices contributing to TI are not needed in order to analyze the marginal deadweight loss of income taxation (as long as individuals behave such that $g_i'(x_i) = \tau$ for all i). However, I argue that knowledge of $dx_i/d\tau$ is useful when designing the income tax system and future tax

[5]Assuming that the individual makes optimal choices for each x_i and that there are no externalities implies that $g_i'(x_i) = \tau$ for all i (Chetty 2009). Thus, based on the envelope theorem, there are no second-order effects on the individual's utility. Originally, the main idea of Chetty (2009) is to show that with weaker assumptions the marginal excess burden is a weighted sum of the total earnings elasticity and the taxable income elasticity. This result holds when the marginal social cost does not equal the tax rate for some x_i, for example, due to fiscal externalities or optimization errors. As highlighted in Chetty (2009), this might be the case in the presence of tax avoidance with transfer costs. Specific theoretical or empirical analysis of this type of framework is, however, out of the scope of this paper.

reforms, even without relaxing the underlying assumptions in the Feldstein model. As pointed out in Blomquist and Selin (2010) and Saez (2003), this information would be valuable if we assume that taxable income itself is directly controlled by the government, which is in fact the case in practical tax policy. The endogenous choice of the tax base is analyzed more thoroughly in Slemrod and Kopczuk (2002) and Kopczuk (2005).

Analysis of the subcomponents of taxable income is more relevant when the assumption of the common income tax rate τ is relaxed. In the extreme case, when different tax rates are applied to all different x_i, equation (4.2.3) can be expressed as

$$DWL = \sum_{i=1}^{n} \tau_i \frac{dx_i}{d\tau_i} \tag{4.2.4}$$

where τ_i represents the tax rate for each x_i.

Abstracting from administrative costs and putting aside tax evasion and tax avoidance, there is no explicit reason to be restricted to a single income tax rate $\tau_i = \tau$ for all of the components of taxable income. Following the assumptions presented so far, in order to minimize the deadweight loss, tax increases should be targeted at choices that are less responsive. On the other hand, the largest economic effects can be achieved when changing the tax rate on the x_i associated with the largest elasticities. In addition to overall ETI estimates, the responsiveness of different types of subcomponents comprising taxable income are in this case the parameters of interest when designing an effective income tax system.

In addition to this Ramsey-type welfare motivation[6], analysis of the anatomy of ETI sheds light on the actual economic nature of the behavioral response. Distinguishing between, for example, real income creation and tax avoidance has important implications for the evaluation of an income tax system (see e.g. Slemrod 1995, 1996, and

[6] In short, the well known Ramsey rule (Ramsey 1927) suggests that goods should be taxed in inverse proportion to their elasticities of demand.

Piketty et al. 2014). Real responses, such as hours of work and work effort, reflect deep individual utility parameters, whereas tax avoidance and tax evasion signal an ineffective and poorly designed tax system. Estimating real and "non-real" subcomponents separately helps to distinguish between the importance of the two in the sense of the marginal excess burden of income taxation.

Finally, a thorough analysis of different subcomponents of taxable income would perhaps call for separate theoretical and empirical frameworks for all of them. However, for the sake of clarity and comparability, I abstract from separate modeling of the different components and approximate them in a single ETI framework, both theoretically and in the empirical model.

4.2.3. Empirical model. This section briefly describes the general empirical methodology of estimating ETI using tax reforms and individual-level panel data.[7] In short, the idea is to measure how the net-of-tax rate affects the taxable income of an individual. Econometrically, this can be described as

$$ln(TI)_{t,i} = e * ln(1-\tau)_{t,i} + ln(\mu)_{t,i} + ln(\lambda)_i + ln(\delta)_t + ln(\varepsilon)_{t,i} \tag{4.2.5}$$

where i denotes the index for individual and t for time. TI is taxable income and $(1-\tau)$ is the net-of-tax rate. $\mu_{t,i}$ denotes other time-variant individual characteristics that affect the income level, and λ_i is a matrix of time-invariant individual characteristics. δ_t is the general time trend and $\varepsilon_{t,i}$ is the individual error term, including the transitory income component.

In practice, it is difficult to identify the average effect of the net-of-tax rate on taxable income using equation (4.2.5). Innate ability and many other time-invariant

[7] For a comprehensive discussion of ETI estimation, including cross-sectional approaches, see Saez et al. (2012). See Saez (2010) and Chetty et al. (2011) for a discussion on identifying ETI locally using the distribution of taxable income and the kink points in the marginal income tax rate schedule.

individual characteristics are unobserved, and at the same time correlated with the progressive tax rate τ. Therefore, in the presence of an income tax reform, one practical approach is to use a first-differences estimator of the form

$$
\begin{aligned}
ln(TI)_{t+k,i} - ln(TI)_{t,i} =& \alpha_t + e(ln(1-\tau)_{t+k,i} - ln(1-\tau)_{t,i}) + \\
& (ln(\mu)_{t+k,i} - ln(\mu)_{t,i}) + (ln(\varepsilon)_{t+k,i} - ln(\varepsilon_{t,i}))
\end{aligned} \tag{4.2.6}
$$

where e is the average elasticity of taxable income. $t+k$ represents the post-reform period, and t represents the pre-reform period. In equation (6), time-invariant individual characteristics are canceled out by definition.

There are many issues that need to be considered before we can achieve a reliable estimate of ETI using equation (4.2.6). These are widely discussed in the empirical ETI literature (see Saez et al. 2012). First, the net-of-tax rate is still endogenous. $(ln(1-\tau)_{t+k,i} - ln(1-\tau)_{t,,i})$ and $(ln(\varepsilon)_{t+k,i} - ln(\varepsilon)_{t,i})$ are mechanically correlated due to the progressive nature of the tax rate schedule (i.e. higher taxable income is taxed at higher marginal tax rates). A positive income shock in year t tends to be followed by lower income in the next period $t+k$, and vice versa. This so-called mean reversion of income combined with the progressive tax rate schedule might bias the elasticity estimate. In addition, non-tax-related changes in the shape of the income distribution need to be taken into account. In particular, if differential variation in tax rates is concentrated only in a certain part of the income distribution, differential income growth trends in different parts of the distribution must be carefully controlled for.

Endogeneity of the net-of-tax rate can be corrected by using instrumental variable estimators. This obviously requires a valid instrumental variable. Non-tax-related changes in income are usually controlled for by adding variants of lagged taxable income and other individual-level controls to the model. Rich individual panel data sets might also allow for controlling the transitory elements of income (see for example Kleven and Schultz 2013). I discuss all of these issues in more detail in Section 4.4.

To summarize, a usual estimable equation for ETI when using individual-level panel data is of the following form:

$$\triangle ln(TI)_{t,i} = \alpha_0 + e\triangle ln(1-\tau)_{t,i} + \alpha_1 ln(B)_{t,i} + \triangle ln(\varepsilon)_{t,i} \tag{4.2.7}$$

where $\triangle$ denotes the difference in the variables between $t+k$ and t, and $(1-\tau)$ is the instrumented net-of-tax rate. In this study, I apply the changes in proportional municipal tax rates as instruments. $B_{t,i}$ is a matrix of individual base-year control variables, including base-year income controls.

4.2.4. Components of total taxable income. In addition to overall taxable income, I also estimate the elasticities of various behavioral choices $\{x_1, ..., x_n\}$ that comprise the overall ETI. The estimable behavioral margins include overall wages, monthly wage rates, fringe benefits, monthly working hours and two specific tax deductions for working individuals, a commuting deduction and a work-related expense deduction. A more detailed description of the components is presented in Table A1 in the Appendix.

The wage rate measures work effort in a broad sense. Separate analysis of fringe benefits examines whether possible effort responses are driven by more irregular and non-monetary components of wages rather than regular cash payments.[8] As a comparison, I estimate the traditional labor supply response in the form of working hours elasticity. This estimate together with the wage rate elasticity sheds light on the extent of real economy responses to income tax rate changes.

[8]Fringe benefit responses can also be considered a type of tax avoidance activity. For example, taxable benefits from the use of a company car are in many cases below the actual opportunity cost of having and driving one's own car. However, the relative advantage of fringe benefits is very case-specific in the Finnish tax system.

The analysis of tax deductions partly reveals the responsiveness of tax planning. A decrease in the net-of-tax rate increases the gains received from decreasing taxable income, and thus increases the incentives to file more deductions than before. Both of the deductions examined in this study are not automatically accounted for in individual taxation. In other words, in order to be eligible for the commuting or expense deductions, a taxpayer needs to fill a tax form and substantiate the desired amount of the deduction.

The list of subcomponents included in the analysis is not exhaustive. This means that I cannot fully construct the total elasticity of taxable income with the (weighted) sum of all the margins estimated in this study. Furthermore, register-based data on hours and wage rates might not be fully reliable, and non-random measurement errors probably occur. I discuss this issue in more detail in Section 4.5.2. Overall, the analysis of the subcomponents is only intended for approximating what the most relevant parts of the behavioral response are in the sense of marginal excess burden.

4.3. Finnish income tax system and recent tax reforms

4.3.1. Institutional setting. I focus on analyzing the behavioral effects of changes in earned income taxation that occurred between 1995-2007. In Finland, earned income is taxed according to a progressive tax rate schedule.[9] In general, the Finnish income tax system follows the principle of individual taxation. The income of a spouse or other family member does not affect the tax rate of an individual. However, some tax deductions and received social security depend on the total income of the household.

[9]Since 1993, Finland has applied the principle of Nordic-type dual income taxation, where earned income (wages, fringe benefits, pensions etc.) and capital income (interest income, capital gains, dividends from listed corporations etc.) are taxed separately. The capital income tax rate is flat. As is typical in a dual income tax system, the top marginal tax rate on earned income (54%) is much higher than the flat tax rate on capital income (28%). Harju and Matikka (2014) present an ETI analysis of capital income and dividend taxation of Finnish business owners.

In Finland there are three levels of earned income taxation: central government (or state-level) income taxes, municipal income taxes, and mandatory social security contributions. All taxes and social security payments are administered centrally by the Finnish Tax Administration.

The central government income tax rate schedule is progressive. The nominal central government income tax rate varies from 0 to 32 per cent[10], depending on taxable income. Social security contributions are proportional. Social security contributions include, for example, mandatory pension contributions and unemployment insurance payments. The average rate of social security contributions is around 5 per cent. Social security contributions are deductible from taxable income. Table A3 in the Appendix presents the schedule for employee social security contributions in 1995-2007.

Municipal income tax rates are flat. The average nominal municipal tax rate is 18.45 per cent. All regular income earners are subject to municipal income taxation, with the exception of individuals with very low earned income who are exempt from all taxes.

There are currently 320 municipalities in Finland (in 2013). Municipalities have autonomous authority to levy income tax. Municipal council elections are held in every four years at the same time throughout the country, and each democratically elected municipal council decides and announces the municipal income tax rate on an annual basis. Figure A1 in the Appendix presents a map of Finnish municipalities and counties in 2007.

There are certain legislative duties and public services each municipality has to offer and fulfill. These include, for example, public health care and social services. These commitments are partly financed by municipal income taxation.[11]

[10]All tax rates presented in this Section are from 2007 if not stated otherwise.

[11]In addition to municipal income tax revenue, the less well-off municipalities receive benefits through local tax-sharing and grants from central government. These are not directly related to the municipal tax rate in the municipality in question. For example, the degree of tax-sharing depends on the

The structure and framework of municipal income taxation, including the flatness of the tax rate and the tax deductions and allowances, are regulated at the central government level. Apart from the need for a certain amount of municipal tax revenue for legislative duties, and the limitations to alter the frame rules of municipal taxation, municipalities can set their income tax rates freely. As a demonstration of this argument, there is a 5 percentage-point difference between the highest (21%) and lowest (16%) municipal income tax rate.

4.3.2. Recent changes in income tax rates.

Central government income taxation. From the mid-1990s onwards, there has been a general decline in central government income tax rates in Finland. Central government tax rates have decreased almost every year in all income classes more or less similarly. Figure 1 illustrates the changes in average marginal tax rates between the years 1995, 2001 and 2007. These marginal tax rates are calculated with the average municipal income tax rate in the year in question. Table A2 in the Appendix presents the marginal tax rate schedule of central government income taxation in 1995-2007.

From the point of view of identification in the empirical ETI model, variation of this sort is not ideal. Although there are significant changes in central government marginal tax rates, the generally declining nature of tax rates does not provide much differential marginal tax rate variation.

Municipal income taxation. Compared to central government income taxes, changes in municipal income tax rates have been different in nature. In Finland, municipal tax rates have changed differently in different municipalities in different years.

industrial and demographic structure of the municipality. Within certain limits, municipalities can also charge usage fees for statutory public services, and assign low real estate taxes. In addition, part of the corporate tax revenue collected by central government is assigned to municipalities.

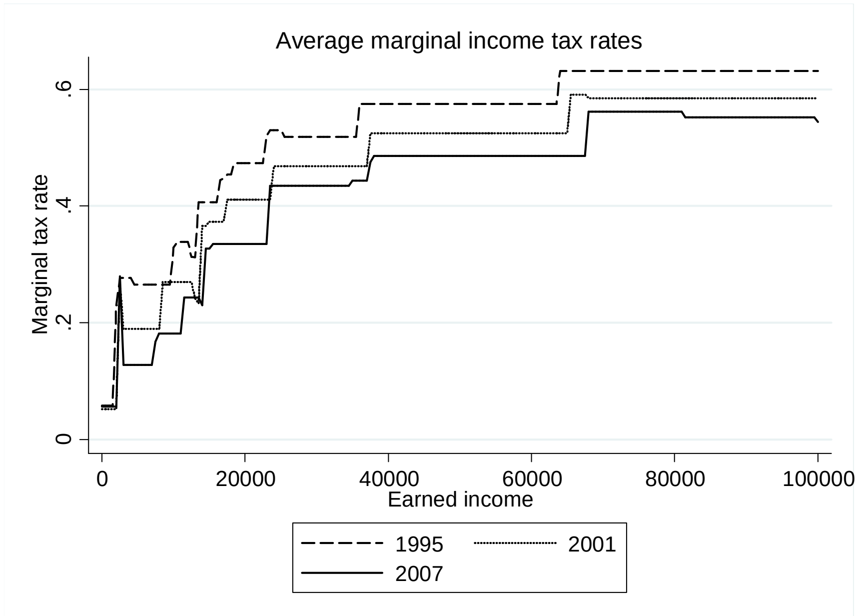


FIGURE 1. Average marginal tax rates in 1995, 2001 and 2007 (calculated with the average municipal tax rate in the year in question)

Table 1 below presents the descriptive statistics of municipal-level tax rate changes in each year. Depending on the year, 10-30 per cent of all municipalities have changed their tax rates. On average, every fifth municipality has changed its tax rate in each year. In all of the years in 1995-2007, at least one municipality has decreased its tax rate, and one has increased it.

Municipal-level tax rate changes vary from -1 to +1.5 percentage points. The average absolute change is approximately 0.5 percentage points. On average, municipal tax rates increased within the time period of 1995-2007. The average municipal income tax rate increased from 17.5% in 1995 to 18.45% in 2007.

There have also been a number of mergers (or consolidations) of two or more neighboring municipalities. Within a merger, the merged municipalities form a new municipality and decide on a new municipal tax rate. As a consequence of mergers, the total number of municipalities decreased from 455 to 416 in 1995-2007. A more detailed discussion on using the *individual-level* municipal income tax rate variation in the empirical analysis is deferred until Section 4.4.2.

Year	Mean absolute change in municipal tax rate (%points)	Std. Dev.	Min change (% points)	Max change (% points)	Percent of municipalities with a change in tax rate	Average municipal income tax rate
1995	0.413	0.166	- 1	0.5	8.9	17.53
1996	0.495	0.220	-1	1	12.0	17.51
1997	0.557	0.207	- 1	1	21.2	17.42
1998	0.548	0.225	- 0.5	1	21.9	17.53
1999	0.558	0.242	- 1	1	21.9	17.60
2000	0.533	0.208	- 1	1	10.3	17.65
2001	0.565	0.219	- 0.5	1.5	25.0	17.67
2002	0.551	0.200	- 0.5	1	20.8	17.78
2003	0.481	0.154	- 0.25	1	11.9	18.04
2004	0.553	0.207	- 0.25	1	31.4	18.12
2005	0.586	0.217	- 0.5	1	31.1	18.29
2006	0.576	0.266	- 0.5	1.5	27.0	18.39
2007	-	-	-	-	-	18.45
Overall	*0.548*	*0.220*	*- 1*	*1.5*	*18.7*	*17.84*

TABLE 1. Municipal income tax rate changes $((t+1)-t)$, 1995-2007

4.4. Data and identification

4.4.1. Data. I use individual-level panel data from 1995-2007, provided by Statistics Finland. The data set consists of approximately 550,000 observations per year, which covers roughly 10% of the Finnish population. The data contains a wide variety of individual-level variables from different statistics. The main statistics used in this study are the personal tax record information provided by the Finnish Tax Administration, the Structure of Earnings statistics collected by Statistics Finland, and available municipal-level statistics.

The data set contains all the necessary information to study the elasticity of taxable income, and a substantial amount of individual and municipal-level control variables. Moreover, the data allow for estimating the tax elasticity of more narrow margins, such as the elasticity of working hours and wage rates based on the Structure of Earnings statistics. Table A5 in the Appendix presents the summary statistics of the key variables used in this study for individuals between 25-60 years of age. Table A5 also includes the descriptive statistics for the key municipal-level variables.

4.4.2. Individual tax rate variation. One of the key issues in identifying ETI is the source of variation in net-of-tax rates. In short, differential variation in net-of-tax rates for otherwise similar individuals is needed when estimating ETI using individual panel data and tax reforms. This study uses changes in municipal income tax rates as the main source of this variation. In the Finnish context, changes in municipal income tax rates are the main source of tax rate variation, as central government income tax rates have decreased rather similarly in all income classes.[12]

Compared to many of the earlier ETI studies, municipal tax rate variation has some very appealing features. First, municipal tax rate changes occur in all of the years in the data (1995-2007). There are also both increases and decreases in municipal tax rates in all of the years.

Second, changes in municipal tax rates affect individuals throughout the income distribution. Thus, in all income classes there are some individuals whose municipal income tax rate has changed, and some individuals faced no changes in municipal income

[12]To my knowledge, Pirttilä and Uusitalo (2005) first proposed the use of municipal income tax rate changes as a source of differential income tax rate variation in Finland.

taxation. This alleviates the potential problems associated with non-tax-related changes in the income distribution, which are critical in many earlier studies.

If the shape of the income distribution varies independently of tax reforms, the analysis of behavioral responses to tax changes might give inaccurate results if this variation cannot be properly taken into account.[13] As changes in municipal income tax rates are not concentrated in certain income classes in any of the years, non-tax-related changes in the income distribution do not bias the elasticity estimates (at least after including appropriate covariates in the model). If nothing else, this bias is certainly smaller than in many of the earlier studies. Furthermore, tax rate variation across the whole income distribution identifies the parameter of main interest, the average elasticity of taxable income.

Figure 2 presents the actual individual marginal income tax rates at different income levels, highlighting the regional variation in marginal income tax rates. As can be seen from this figure, individuals with the same level of income face different marginal tax rates depending on the municipality of residence. Moreover, with regard to identification, individuals with the same income level face different *changes* in overall marginal tax rates due to differential changes in municipal tax rates over time.

Table 2 describes the overall individual variation in municipal income tax rates. In addition to tax rate changes within a municipality, Table 2 includes individuals who faced a change in their municipal tax rate as a result of a change in their municipality of residence, or as a consequence of merger of two or more neighboring municipalities. In the data set, 3.3% of individuals changed their municipality of residence between t and $t+1$ (on average). This number does not include mergers of municipalities. I

[13]In Finland, the overall income distribution somewhat polarized between 1995-2007 (see Riihelä, Sullström and Suoniemi 2008). However, changes in the distribution are mostly driven by changes in capital income, not by changes in earned income, to which I focus on in this study. Changes in the income distribution are also relatively modest compared to, for example, the US in the 1980s.

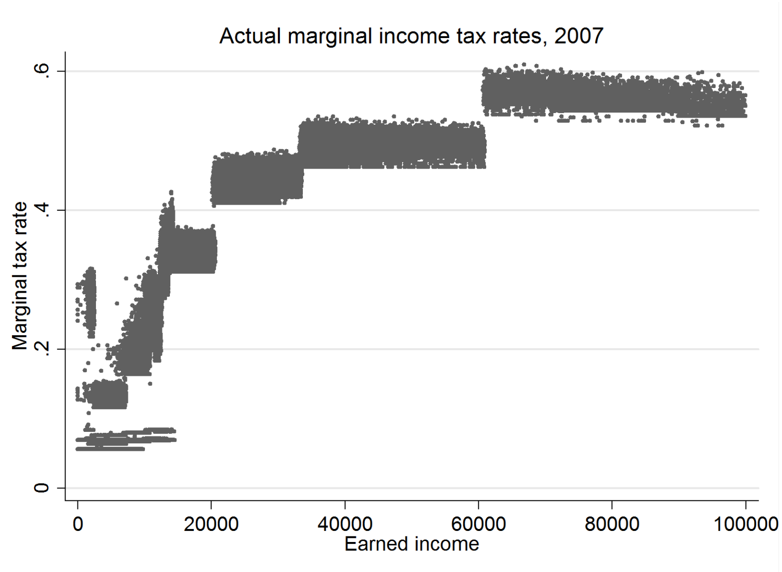


FIGURE 2. Actual marginal tax rates in 2007, including individual municipal income tax rates

discuss the implications of individuals changing their municipality of residence in the next subsection.

As can be seen from Table 2, approximately every fifth individual experienced a change in his/her municipal income tax rate each year. On average, the absolute change in the municipal tax rate was 0.6 percentage points for those individuals who faced a change in their municipal tax rate. There is a more distinctive difference between the smallest negative (-4.25 percentage points) and largest positive (3.75 percentage points) change in the municipal tax rate. The largest absolute changes are caused by changes in the municipality of residence.

Individual changes in municipal income tax rates are not very large in size. The majority of changes are between +/- 0.25-1 percentage points. When the whole net-of-tax rate is accounted for (municipal taxes + central government taxes + social security contributions), most of the changes are around +/- 1-10 as a percentage. The largest

Year	Mean absolute change in municipal tax rate (%points)	Std. Dev.	min change of munic. tax rate (% points)	max change of munic. tax rate (% points)	Percent of individuals with a change in municipal tax rate
1995	0.533	0.331	-3.25	3.75	9.6
1996	0.508	0.250	-3.25	3.5	22.2
1997	0.632	0.272	-3	2.75	24.2
1998	0.601	0.289	-3	3.5	20.4
1999	0.564	0.307	-3.25	3.75	17.5
2000	0.608	0.341	-3.75	3.5	11.7
2001	0.605	0.291	-3.25	3.25	23.4
2002	0.716	0.301	-3	3.5	30.6
2003	0.581	0.243	-2.75	3.0	17.7
2004	0.634	0.246	-3.5	3.25	29.7
2005	0.596	0.260	-3.5	3	22.2
2006	0.599	.0316	-4.25	3.75	15.2
Overall	*0.608*	*0.288*	*-4.25*	*3.75*	*18.7*

TABLE 2. Individual-level tax rate variation $((t+1)-t)$, 1995-2007

changes in municipal tax rates correspond to changes in overall net-of-tax rates of +/- 5-15%.

Some recent studies (e.g. Chetty 2012, Chetty et al. 2011, Kleven and Schultz 2013) underline that optimization frictions have an effect on the estimated taxable income elasticity. In short, if costs related to responding to tax changes (adjustment costs, job search costs, paying attention to tax code, filing deductions etc.) are large, they might attenuate the observed elasticities and make them less than the structural elasticities derived in a frictionless benchmark case.

In general, most of the frictions are more relevant when changes in the tax rate schedule are small. Small tax rate changes might induce only small utility benefits from changing behavior, and this utility gain might be smaller than the associated (fixed) costs. Thus small changes in tax rates tend to lead to smaller changes in observed behavior (on average).[14] This is a valid point in this setup, as the variation in overall net-of-tax rates is relatively small, at least when compared to many earlier studies. Therefore, assuming that adjustment costs or other frictions matter, we would expect to get smaller ETI estimates in this study.

On the other hand, small tax rate changes have high policy relevance. Usually income tax reforms are not particularly large. Many recent reforms in industrialized countries can be regarded more or less as fine-tuning of the tax systems. Therefore, a careful study of smaller-scale tax reforms might have greater practical relevance than analysis of more extensive and unique reforms, such as the tax rate cut of 1986 in the US.

In addition, it might be that the short-run response to a small change in the net-of-tax rate differs significantly from the longer-run effect, especially in the case of adjustment or search costs. Adjustment to a new level of income tax rate might take more than 1-3 years, particularly if the short-run gains from the behavioral response are relatively small. In the empirical part, I also test the effect of changing the time horizon in the elasticity estimate.[15]

Finally, as highlighted by Kopczuk (2005), changes in the tax base and the definition of taxable income probably affect the ETI estimate. In Finland, the definition of taxable earned income has remained relatively constant between 1995-2007. Furthermore, the minor changes in the tax base are, at least to some extent, unrelated to the main source of differential tax rate variation. This is due to the fact that the tax base and basic rules

[14] Using Danish data, Chetty et al. (2011) and Kleven and Schultz (2013) show evidence that the observed elasticity estimate depends positively on the size of the change in the net-of-tax rate.

[15] However, as noted in Gruber and Saez (2002), theoretical prediction of the effect of the time window on the elasticity estimate is not clear. It might also be that individuals react to tax rate changes actively in the short run, and then return to their original level of taxable income in the longer run (see for example Goolsbee 2000). Gruber and Saez (2002) find no significant time horizon effects in their study. In contrast, Giertz (2010) reports that elasticity tends to increase as the time horizon increases.

of municipal income taxation, including tax deductions and allowances, are regulated at the central government level.

4.4.3. Net-of-tax rate instrument. In a progressive income tax rate schedule, the marginal income tax rate increases as taxable income increases. Therefore, a change in taxable income endogenously defines the change in the net-of-tax rate. Thus the elasticity coefficient in equation (4.2.6) is very unlikely to capture the actual behavioral response to a tax rate change without using an instrumental variable estimator, and therefore a valid instrumental variable for $(1-\tau)$ is required.

A common strategy in the earlier literature is to simulate predicted or synthetic net-of-tax rates, and use them as instruments for the actual net-of-tax rate changes (see for example Gruber and Saez 2002). The basic structure of a predicted net-of-tax rate variable is the following: take base year t income and use it to predict the net-of-tax rates for $t+k$ by using the post-reform tax legislation in $t+k$. The synthetic net-of-tax rate instrument is then the difference between the actual net-of-tax rate in t and the net-of-tax rate calculated with income in t and the tax law for $t+k$. The intuition is that the predicted difference describes the exogenous change in tax liability caused by changes in tax legislation, while ignoring any behavioral effects by keeping taxable income constant.

However, the predicted net-of-tax rate variable is a function of individual taxable income in year t, and there is no proof that this instrument is exogenous in the empirical model. Following Blomquist and Selin (2010) and Moffit and Wilhelm (2000), it is unlikely that the predicted net-of-tax rate instrument is correlated similarly with both $\varepsilon_{t+k,i}$ and $\varepsilon_{t,i}$ in equation (4.2.7), as taxable income in year t defines the marginal tax rate in both t and $t+k$. In addition, there is no general proof that the usually added controls, mainly base-year taxable income and other individual characteristics, correct this endogeneity problem (see Weber 2014). All in all, there is concern about the validity of instruments that are explicit functions of the dependent variable.[16]

In this study I use an instrument for the net-of-tax rate changes which is not a function of taxable income, namely changes in proportional municipal income tax rates. As the municipal income tax rate is flat, the tax rate is the same in all income classes within each municipality. In other words, at the individual level, the only determinant of the municipal income tax rate is the municipality of residence.[17]

Compared to previous studies, I do not have to make assumptions about the time structure of the individual transitory income component in order to ensure the exogeneity of the instrument. In addition, as municipal income tax rates affect the net-of-tax rates in all income classes, I do not have to explicitly control for the non-tax-related changes in the income distribution in order to guarantee the causality of the behavioral parameter. Furthermore, mean reversion does not pose a notable problem, as yearly fluctuation in individual income does not affect the instrument.

Even though the municipal tax rate instrument is not a direct function of the dependent variable in any period, there are some concerns that the instrument is not exogenous as such. The main concern is the possible policy endogeneity of municipal tax rate changes. In other words, municipal tax rate changes are probably not randomly assigned in the population.

[16]In Blomquist and Selin (2010), the middle-year characteristics (i.e. $(t+t+k)/2$) are used to define imputed taxable income for both t and $t+k$, from which the net-of-tax rate instrument is then calculated. Blomquist and Selin (2010) show that this strategy produces exogenous instruments under relatively general assumptions about the autoregressive structure of the transitory income component. Weber (2014) shows that using lagged income ($t-1$, $t-2$ etc.) to derive the instrument enhances the validity of the predicted instrument. Nevertheless, the validity of these types of predicted net-of-tax rate instruments is still dependent on the serial correlation pattern of $\varepsilon_{t,i}$.

[17]The earned income tax allowance in municipal taxation depends (inversely) on earned income. This mainly affects low-income individuals. The effect of the allowance on the effective overall net-of-tax rate is trivial for taxable income over 14,000 euros. The earned income tax allowance in municipal taxation is described in detail in Table A4 in the Appendix.

In order to alleviate potential policy endogeneity, I include various municipal-level covariates to the model, such as municipal-level unemployment and employment rates, net migration and the level of net debt (per capita). All of these variables have a presumable effect on total taxable income within a municipality, as well as average individual taxable income. For example, municipalities might increase tax rates when future tax revenue losses are predicted. This can be caused by decreased employment in the jurisdiction. Because low employment might also decrease individual taxable income (on average), the elasticity estimate may be upward-biased. By including a set of municipal-level covariates and regional income trends in the model, I can, at least to a reasonable extent, separate the possible municipal and other regional-level effects from the individual-level behavioral responses. Furthermore, in Section 4.4.4, I test the supposition that future tax increases (or decreases) are more common when taxable income has decreased (Figure 4).

Another cause for concern is the possibility that individuals select into the "treatment" by changing their municipality of residence. First, we might worry that individuals consistently move to municipalities with lower (or higher) tax rates. However, with regard to identification in the ETI model, this is not very relevant in itself.[18]

A more serious concern would be that changes in taxable income are systematically correlated with the moving decision, and especially with the municipal tax rate in the destination municipality (i.e. the tax rate instrument is correlated with the transitory

[18]For example, if an individual moves to a municipality with a lower tax rate but does not change his/her current job (or more precisely, taxable income does not change), ETI for this individual will be zero by definition, even though the total taxes paid are now lower than before. Thus, this kind of purely tax-motivated migration is not an issue in this framework. Also, we might suspect that there is a classical selection problem in equation (4.2.5). The conceivable selection bias comes from the possibility that individuals who prefer low income taxation choose to reside in a municipality with a low tax rate. Preference for low income taxation is likely to be positively correlated with taxable income, causing the elasticity estimate to be biased. However, as the empirical model in question is identified by individual changes in both municipal tax rates and taxable income, this is not a very serious concern.

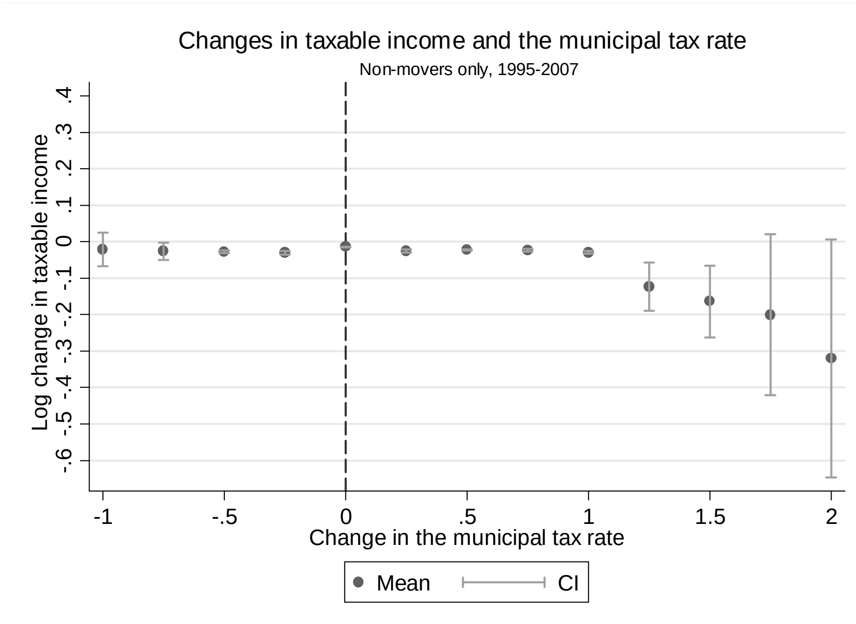
income component). For example, a new, better paid job might be a good reason for moving to another municipality. At the same time, it could be that municipalities with a lot of open highly paid vacancies have a relatively low or high municipal tax rate, which would cause bias in the elasticity estimate. This is a relevant concern in the Finnish case, as the municipal income tax rates are below the average in high-wage regions such as the capital city area (Helsinki-Espoo-Vantaa).

Therefore, in the preferred empirical specification, I drop individuals who change their municipality of residence between t and $t+k$ in order to avoid mechanical correlation between the instrument and the transitory income component.[19] The downside of this approach is that it does not take into account the potential (but somewhat unlikely) effect of moving and decreasing (increasing) income because of a municipal tax increase (decrease) in the base-year municipality of residence. In other words, the model without changes in the municipality of residence does not include all potential behavioral margins. I further discuss this issue in Section 4.5.3.

4.4.4. Descriptive statistics. Figure 3 describes the connection between changes in individual taxable income and changes in municipal tax rates. In the Figure, I plot mean changes in log taxable income by different changes in the municipal tax rate between $t+1$ and t. Plotting mean changes in taxable income by changes in municipal tax rates is feasible as changes in municipal tax rates occur in 0.25 percentage point intervals (0.25, 0.5, 0.75 etc.). For example, the point on the vertical dash-line in the Figure denotes the average log change in taxable income between $t+1$ and t for those individuals who faced no changes in their municipal tax rate in the same time period.

From Figure 3 we can see that relative changes in taxable income are, on average, more negative the larger the positive changes in municipal tax rates are. In other words,

[19] In order to test the effect of moving individuals, I also estimate the model with the movers included. In this case, I add an individual moving dummy to the estimable equation, along with the interaction terms of the moving dummy and the destination county. This controls for the average effect of moving to a certain region on individual income (given other individual characteristics).

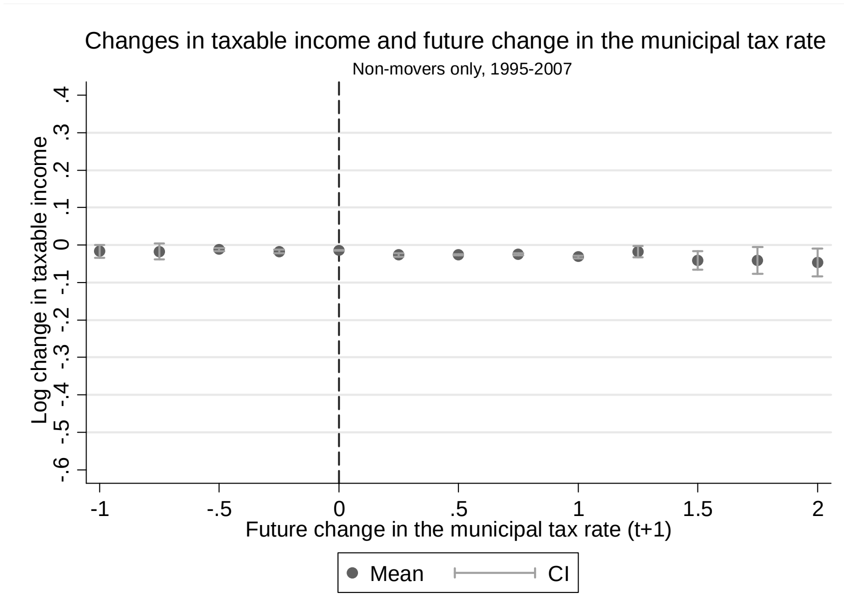


Notes: The baseline sample includes observations where base-year taxable income is above 20,000 euros. Pensioners and people under the age of 24 and over the age of 60 are not included in the sample. Also, the sample is limited to individuals whose absolute change in log taxable income between t and $t+1$ is below 8.5, and whose marital status is unchanged between the two years. For more details, see Section 4.4.5.

FIGURE 3. Changes in taxable income and changes in municipal tax rates

positive changes in municipal tax rates induce negative changes in taxable income. This reduced-form type evidence suggests that individuals respond to incentives created by changes in municipal tax rates. In addition, Figure 3 graphically illustrates the extent of municipal tax rate variation for the baseline estimation sample included in the Figure.

Figure 4 shows the mean changes in log taxable income with respect to *future* changes in the municipal tax rate, i.e. changes in the municipal tax rate between $t+2$ and $t+1$. Intuitively, if municipalities respond to a decrease in taxable income in the past by increasing the municipal tax rate in the future, we should see that future tax increases are more common when there is a decreasing trend in average taxable income (and vice versa).



Notes: The baseline sample includes observations where base-year taxable income is above 20,000 euros. Pensioners and people under the age of 24 and over the age of 60 are not included in the sample. Also, the sample is limited to individuals whose absolute change in log taxable income between t and $t+1$ is below 8.5, and whose marital status is unchanged between the two years. For more details, see Section 4.4.5.

FIGURE 4. Changes in taxable income and future changes in municipal tax rates

Figure 4 does not support this policy endogeneity channel. There is no statistical difference between the changes in taxable income with respect to future changes in municipal tax rates of different size, which suggests that future tax changes are not determined by past changes in individual taxable income. Nevertheless, in order to take into account possible policy endogeneity, I add municipal-level covariates and other regional controls in the preferred empirical specification.

Figures 3 and 4 include the baseline estimation sample where individuals who change their municipality of residence between t and $t+1$ are dropped out. Figure A2 in the Appendix shows a similar picture for the sample including the movers. The Figure including the movers delivers similar conclusions as before. The left-hand side of Figure

A2 shows that tax increases lead to a negative change in mean taxable income, and vice versa. From the right-hand side of Figure A2 we can see that endogeneity based on past changes in average taxable income is not the driving force behind the results.

4.4.5. Estimable equation. I estimate different variations of the following equation using a two-stage least squares estimator

$$\begin{aligned}\triangle ln(TI)_{t,i} =& \alpha_0 + e\triangle ln(1-\tau)_{t,i} + \alpha_1 f(lnTI)_{t,i} + \\ & \alpha_2 B_{t,i} + \alpha_3 M_{t,m} + \sum_j \alpha_{4j} YEAR_j + \triangle\varepsilon_{t,i}\end{aligned} \tag{4.4.1}$$

In equation (4.4.1), $\triangle ln(TI)_{t,i}$ is the change in taxable income between t and $t+k$ for individual i. $\triangle ln(1-\tau)_{t,i}$ is the change in the net-of-tax rate instrumented with the change in the municipal net-of-tax rate. Thus e is the coefficient of interest, the average elasticity of taxable income with respect to the net-of-tax rate.

Despite the fact that in this setup the non-tax-related changes in the income distribution and mean reversion are not as problematic as in many earlier studies, I add a ten-piece base-year taxable income spline variable (denoted by $f(lnTI)_{t,i}$) into the model in some specifications. This income control serves as a proxy for individual unobserved heterogeneity in income growth (Blomquist and Selin 2010).

Base-year variables $B_{t,i}$ control for observed individual heterogeneity affecting changes in taxable income. $B_{t,i}$ includes age, age squared, county of residence, sex, level of education (highest degree), marital status, size of the household and dummy variables indicating whether the individual has received any taxable social security benefits in the base year.[20] I also include interaction terms of sex and other controls in the model

[20]Social security benefits include unemployment benefits, sickness benefits, parental leave benefits and study grants.The marital status dummies include married couples, unmarried couples, singles, divorced singles and widows/widowers. There are 21 counties in Finland. Figure A1 in the Appendix presents a map of the counties.

(age, education, household size and marital status). Importantly, I also add county-year fixed effects, which control for different income trends in different parts of the country at different times.

To control for the possible policy endogeneity of the net-of-tax rate instrument, I add municipal-level (m) characteristics $M_{t,m}$ to the estimable equation in some specifications. $M_{t,m}$ includes base-year values of municipal-level employment, unemployment, net migration and net loan positions. These variables reflect the actual publicly available information that the decision-making bodies in each municipality have on the local economy. Finally, I add year dummies to control for time.

I limit the analysis to observations where base-year taxable income is above 20,000 euros. First, the income cut-off is needed in order to eliminate any notable effect of the municipal earned income tax allowance on the net-of-tax rate instrument. Second, I focus on the intensive margin behavioral responses, which emphasizes the need for an income cut-off. Many of the social security benefits in Finland (e.g. unemployment benefits and sickness benefits) are regarded as taxable income, which creates relatively low but positive taxable income also for individuals fully or partly outside the labor force.

In addition, I drop pensioners and people under the age of 24 and over the age of 60 out of the estimation sample. Also, following earlier literature, the analysis is limited to individuals whose absolute change in log taxable income between t and $t+k$ is below 8.5, and whose marital status is unchanged between the two years. Finally, in the baseline analysis, I drop individuals who change their municipality of residence between t and $t+k$. However, the sample includes individuals whose municipality of residence changed due to a municipality merger.[21]

[21] As a sensitivity check, I also estimate the model with movers included. For the models including the movers, $B_{t,i}$ also contains a dummy variable denoting whether an individual has changed his/her municipality of residence between t and $t+k$, and the interaction terms of the moving dummy and the county of residence in $t+k$.

The baseline time horizon used is three years, which is customary in the literature. In order to be able to separate this middle-term elasticity from the shorter-run effects, I drop all the observations where the individual municipal income tax rate also changed between $t+1$ and $t+2$, or $t+2$ and $t+3$. Finally, as sensitivity checks, one and five-year difference models are also estimated along with other alternative specifications.

Equation (4.4.1) is also used to estimate the subcomponents of overall taxable income. The subcomponents include overall wage income, monthly wage rates, taxable fringe benefits, hours of work and two itemized work-related tax deductions (commuting cost and work-related expense allowances). The same set of controls and sample limitations are applied in the estimation of these margins.

4.5. Results

4.5.1. Taxable income elasticity. Table 3 shows the results for the three-year difference model with different specifications.[22]

First, column (1) shows the ETI estimate with only year fixed effects included in the regression. This estimate is approximately 0.6 and statistically significant at the 1% level. Adding controls in columns (2)-(4) decreases the point estimate. In column (2), the point estimate is 0.38 and statistically significant when the 10-piece base-year income spline variable and individual base-year controls are included.

In column (3) I do not include base-year income splines in the equation. Without income splines the estimate is close to that with the splines included (0.33). First, this implies that income controlling does not have much effect on the ETI estimate in this particular case in which the net-of-tax rate instrument is unrelated to individual income. This can also be seen as tentative evidence that non-tax-related changes in the income

[22]The F-statistics for the first stage of the two-stage least squares routine are very large and highly significant in all specifications. The first-stage result for the baseline specification is presented in Table A6 in the Appendix. The standard errors are clustered at the municipal level in every specification. Clustering is needed because the error terms might be correlated between individuals residing in the same municipality. However, clustering has only a small increasing effect on the standard errors.

VARIABLES	(1) TI	(2) TI	(3) TI	(4) TI
Elasticity	0.618***	0.383***	0.332**	0.266**
	(0.227)	(0.139)	(0.151)	(0.132)
Year F.E.	Yes	Yes	Yes	Yes
County F.E.	No	Yes	Yes	Yes
County-Year F.E.	No	No	No	Yes
Base-year income spline	No	Yes	No	No
Other base-year controls	No	Yes	Yes	Yes
Municipal-level controls	No	No	No	Yes
Observations	414,645	413,862	414,221	413,482

Robust and municipal-level clustered standard errors in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

TABLE 3. ETI estimates

distribution do not notably affect the elasticity estimates when tax rate variation occurs at all income levels. Second, base-year income is not an exogenous variable in the first-differences setup, and thus not an optimal choice as a control variable. Therefore, there are no explicit reasons why these variables need to be added to the ETI model in this case, and thus I prefer a specification in which income splines are not included.

Column (4) in Table 3 shows the preferred empirical specification with extensive regional controlling. First, I add the interactions of county and year fixed effects to control for different income trends in different years in different regions. Second, I add base-year municipal-level variables. As mentioned before, there are reasons to suspect that municipal tax rate variation is not randomly assigned across individuals in different municipalities (given other individual characteristics). Therefore, controlling for municipal-level characteristics $M_{m,t}$ might be needed in order to alleviate this potential policy endogeneity.

After adding county-year fixed effects and municipal controls, the ETI estimate is 0.27 and statistically significant at the 5% level. This estimate is broadly in line with many previous ETI studies, although it is larger than the average ETI in most recent papers from other Nordic countries (Kleven and Schultz 2013, Thoresen and Vattø 2013). One of the reasons for the larger point estimate might be the different identification strategy. Instead of using the predicted net-of-tax rate, I use changes in flat municipal tax rates as the net-of-tax rate instruments. I further discuss this issue in Section 4.5.3.

4.5.2. Subcomponents of taxable income. The results for subcomponents of overall taxable income are presented in Table 4. All of the models include year, county and county-year fixed effects, individual base-year controls and municipal controls.

VARIABLES	(1) Wage income	(2) Monthly wage	(3) Monthly hours	(4) Fringe benefits	(5) Work related expenses	(6) Commuting expenses
Elasticity	0.485*	-0.03	0.178	1.00	-0.298**	-0.923
	(0.247)	(0.133)	(0.138)	(1.301)	(0.138)	(1.919)
Observations	399,417	244,116	242,412	121,576	398,324	123,919

Robust and municipal-level clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
All regressions include year, county and county-year fixed effects, individual base-year controls and municipal controls.

TABLE 4. Elasticity estimates for subcomponents of taxable income

First, column (1) shows the elasticity estimate for the overall yearly wage income. The wage income information comes from the Finnish Tax Administration.[23] Yearly wage income includes fringe benefits and other irregular earnings categorized as compensation for working. The elasticity of wage income is relatively large (0.49), which implies that wage income as a whole responds to the income tax rate.

[23]The separation of wage income and other earned income is important in the Finnish tax system. For example, some tax deductions are only based on wage income, and not other types of earned income such as taxable social benefits.

Columns (2) and (3) show the elasticity estimates for monthly wage rates and monthly hours. The information on monthly wages and monthly hours comes from the Structure of Earnings statistics collected by Statistics Finland. Monthly wage rates include both regular and irregular earnings as well as fringe benefits. Monthly working hours include regular hours and overtime working hours.

The point estimate for the wage rate is close to zero (-0.03), and insignificant. However, monthly hours appear to be moderately responsive, although the estimate (0.18) is not statistically significant. These estimates imply that both work effort and labor supply are not very responsive to income taxation, suggesting that the real-term responses are small or moderate at the most.

There is a slight conflict between the estimates in column (1) and columns (2) and (3). It seems that while yearly wage income is relatively responsive to taxes, the wage rate and working hours are not. There are plausible data-driven explanations for this finding. First, information on monthly wages and monthly hours are only collected for a selected sample of full-time workers in companies with more than five workers. This might affect the results if part-time workers or workers in smaller firms can respond more flexibly to tax incentives. This assumption is also supported by the data. The ETI estimate for the subgroup with non-missing monthly hours is 0.22 (0.11), which is lower than the baseline estimate in column (4) of Table 3 above. This indicates that individuals included in the Structure of Earnings statistics have, on average, somewhat lower responsiveness to tax incentives.

Furthermore, wage rates and working hours are mainly based on the situation in October in each year, which might not reflect the actual yearly responses, especially with respect to more irregular components such as fringe benefits or overtime hours. In addition, working hours and wage rates are reported by employers, and thus they might not precisely measure the actual wage rates or working hours of each individual worker, especially if wages are not directly based on actual hours worked (i.e. workers with a fixed monthly salary with no overtime compensations). Nevertheless, given the limitations of the Structure of Earnings data, I find no evidence of extensive and significant effort or labor supply responses to tax changes. Based on this evidence, it seems that more irregular and flexible components of taxable income and total wage income might drive the results.

To further study the potential effects of other components, I estimate separate elasticities for fringe benefits and two specific tax deductions for working individuals, the work-related expenses deduction (tulonhankkimisvähennys) and commuting expense deduction (työmatkavähennys). Both these deductions are not automatically accounted for in personal income taxation, and need to be itemized by the taxpayer in order to qualify for the deduction (for more details, see Table A1 in the Appendix). The data on taxable fringe benefits and tax deductions come from the Finnish Tax Administration.

Column (4) of Table 4 presents the elasticity estimate for taxable fringe benefits. The responsiveness of fringe benefits seems to be relatively large (elasticity of 1), although the effect is imprecisely measured. This evidence tentatively supports the view that the response might come through more irregular earnings channels.

Columns (5) and (6) show the estimates for the deductions. Both deductions seem to be rather responsive to tax rate changes, and the elasticity of work-related expenses is statistically significant at the 5% level. In contrast, the commuting deduction response is very imprecisely measured (mostly due to the relatively small number of observations in the data). The signs of both deduction responses are intuitive, however. Basic taxable income theory predicts that the amount of tax deductions will increase as the net-of-tax rate decreases, and vice versa. This evidence together with the relatively large fringe benefit response tentatively imply that the overall ETI is driven by tax deduction behavior and irregular earnings rather than conspicuous changes in labor supply or work effort.

4.5.3. Alternative specifications and sensitivity checks. Table A6 in the Appendix presents the results for alternative specifications and sensitivity checks for the average ETI model. First, column (1) in Table A6 shows the estimate for the baseline specification including individuals who move from one municipality to another between t and $t+3$. The ETI estimate increases to over 0.6 when movers are included in the sample. This implies notable effects for individuals who changed their municipality of residence. As mentioned in Section 4.4.3, the elasticity estimate might be larger for movers due to possible mechanical correlation between the instrument and transitory income that cannot be fully taken into account with the available covariates.[24]

On the other hand, the smaller point estimates in Table 3 might also indicate that costs and benefits related to optimization behavior matter, and individuals do not respond as actively to smaller changes in marginal tax rates. Individuals who move face larger changes in their net-of-tax rates (on average), which provides greater incentives to alter their behavior as well.

Column (2) shows the elasticity estimate for gross earned income subject to taxation, which is a broader income concept than taxable income. The point estimate for gross earned income is slightly lower (0.20) than for taxable income. This is an expected result because, for example, taxable income is subject to more deductions than gross earned income. Column (3) shows the elasticity estimate with income weights. The

[24]As a further robustness check, I also estimate the model without individuals who move to the largest county in Finland, which includes the capital city area (Uusimaa). The average point estimate for this model is 0.43, which is lower than the estimate when all movers are included. This supports the view that mechanical correlation between the lower-than-average municipal tax rates and higher wage levels in larger cities might bias the results when moving individuals are included to the model. Second, I estimate the preferred model including the movers by using a municipal tax rate instrument based on only the tax rate changes in the base-year municipality of residence. In this model, the tax rate of the destination municipality is not taken into account for movers, as the tax rate instrument is not affected by the moving decision. This model gives a similar ETI estimate (0.29 (0.18)) as the baseline model. This supports the view that the potential effect of both moving and altering taxable income due to changes in the tax rate of the base-year municipality of residence is not very relevant.

income-weighted point estimate (0.27) does not significantly differ from the unweighted baseline estimate.

Columns (4) and (5) present the estimates separately for men and women. The results show that the point estimate for men (0.23) is smaller than for women (0.43). Columns (6)-(8) present the estimates for three different income levels: low income (10k-25k euros), middle income (25k-40k euros) and high income (over 40k euros). The results show that low-income (point estimate 0.33) and high-income (0.97) individuals seem to be more responsive to income taxes than middle-income individuals (0.13). This tentatively suggests that the elasticity follows a U-shaped curve in which low and high-income individuals have the largest elasticities.

Column (9) shows the baseline estimate without including individuals in the capital city area (Helsinki, Vantaa, Espoo and Kauniainen). Average income levels and income growth are higher in the capital city area, and it also might be that individuals in this area respond differently to tax changes. However, dropping the capital city area form the estimation sample does not significantly affect the point estimate (0.25).

Columns (10)-(12) study different time horizons. Column (10) presents the ETI estimate for the one-year difference model (0.43), which is slightly larger than the baseline three-year estimate. Column (11) shows the estimate for the five-year model. In general, the five-year model produces very imprecise results, the point estimate being -0.1. The reason for this is the chosen identification strategy where I have dropped all individuals whose municipal income tax rate has also changed between $t+1$ and $t+2$, or $t+2$ and $t+3$ in the baseline three-year model. Extending this condition to the five-year setup substantially reduces the number of observations available, and produces a selective sample of individuals in municipalities with only a few tax rate changes in 1995-2007.

Alternatively, in order to estimate a longer-run response, I pool two six-year differences (2001-1996 and 2007-2002) together and estimate the model without dropping individuals with changes in municipal tax rates in the middle of the differences. Column (12) shows that the point estimate for this regression is 0.67, and statistically significant at the 5% level. Thus it seems that the longer-run response is larger than the short-run, although the point estimate from this particular model is a mixture of both long and short-run responses.

Column (13) shows the estimate when using the standard Gruber and Saez (2002) type predicted net-of-tax rate instrument discussed above in Section 4.4.3. Similarly as in Blomquist and Selin (2010), I get a negative point estimate when applying this instrument (-0.23). This implies that there might be a bias in the ETI estimate when using the predicted net-of-tax rate instrument. This is especially plausible in the Finnish setting. For example, on average, net-of-tax rates increased in all income classes in 1995-2007. However, the relative growth in net-of-tax rates was larger for middle and low-income earners compared to high-income earners (see for example Table A2 in the Appendix). Thus we might get biased estimates if high-income earners have even slightly faster non-tax-related income growth which we cannot fully control for with the available controls.

This is also supported by the data. Using the same income classification as above, the point estimates for low and middle-income are close to zero (-0.057 and -0.022, respectively) when using the predicted tax rate instrument. In contrast, the estimate for high-income individuals is -0.61. This indicates that the negative average effect of the predicted tax rate instrument is driven by high-income earners in the top income brackets.

I further study the predicted instrument by omitting the municipal tax rate variation when constructing the predicted tax rate instrument. In other words, I apply average municipal tax rates in each year, and use only the variation stemming from recurrent decreases in the progressive central government earned income tax rates (see Table A2 in the Appendix) when deriving the instrument. Results show that omitting the variation in flat tax rates further decreases the average point estimate with the predicted instrument (-0.38). Point estimates also decrease in different income classes (-0.19, -0.15 and -0.76 for low, middle and high-income earners, respectively). To summarize, these observations confirm the results from earlier studies (Weber 2014, Blomquist and Selin 2010) which highlight the potential inconsistency of predicted net-of-tax rate instruments derived using base-year taxable income and variation stemming from different tax rate changes in different income groups.

Finally, columns (14)-(16) show the OLS, first-stage and reduced-form results, respectively. First, column (14) shows that the OLS estimate for the ETI model provides a highly counterintuitive result with a large negative point estimate (-2.9), which highlights the need for a valid instrumental variable. Column (15) shows that the first-stage results are strong. The first-stage estimate implies that a 1% increase in the municipal tax rate accounts for an approximately 1.4% increase in the overall net-of-tax rate. Given the general pattern of the central government tax rate changes and the three-year time window, this estimate is reasonable in size. Also, as mentioned before, the F-statistics for the first-stage models are large and highly significant in all specifications (824 in column (15)). Column (16) shows the results for the reduced-form model where the log change in taxable income is regressed directly with the log change in the net-of-municipal tax rate. The results show that, on average, individuals respond to municipal tax rate changes actively and in a statistically significant manner, which is also illustrated above in Figure 3.

4.6. Conclusions

In this study I analyze the key tax policy parameter, the elasticity of taxable income (ETI), using Finnish panel data from 1995-2007. In addition to overall ETI estimates, I outline the responsiveness of various subcomponents of taxable income, such as labor supply and tax deductions. This anatomy of the overall responses is rarely analyzed in the ETI literature.

I use variation in flat municipal income tax rates as an instrument for the changes in overall net-of-tax rates. The flat municipal income tax rate is not a function of individual taxable income in any period. Also, changes in municipal income tax rates occur in all income classes in all years, and in both directions. Therefore, non-tax-related changes in the shape of the income distribution and mean reversion are not problematic, as the instrument is unrelated to the level of individual income, and the net-of-tax rates vary differently throughout the income distribution. The novel approach of using changes in municipal tax rates as instruments underlines that different institutional features can provide useful and practical variation in terms of estimating sufficient statistics for welfare analysis.

My preferred estimate for the average ETI in Finland is 0.27. This estimate is in line with many previous studies from other countries. Interestingly, the estimate is somewhat larger than recent estimates from other Nordic countries (see for example Kleven and Schultz 2013). This might be partly due to different estimation strategies. It is possible that the net-of-tax rate instruments used in previous studies provide estimates that are more or less biased.

The ETI of 0.27 suggests that the welfare loss of income taxation is moderate in Finland. At the average point, the ETI estimate implies a marginal excess burden of around 15% (see Section 4.2.2), which is in line with many previous ETI studies. Intuitively, this implies that increasing income taxes in Finland would induce non-negligible but not extensive efficiency losses.

The subcomponent analysis suggests that real behavioral margins, such as working hours and wage rates, respond less than tax deductions and more irregular forms of compensation such as fringe benefits. The results show no extensive or statistically significant responses to register-based monthly hours and monthly wages, whereas the point estimates for tax deductions and fringe benefits are large, although mostly imprecisely measured. This tentatively implies that the overall behavioral response is not driven by profound economic parameters such as the opportunity cost of working. Thus even though the average ETI estimate is not trivial, changing the income tax rate seems to have only a limited effect on labor supply and work effort, especially for full-time workers in larger firms.

However, the results from the subcomponent analysis need to be interpreted with some caution. It is possible that register-based data on working hours and wage rates are not sufficient to adequately measure real behavioral margins. Thus, in future work, we need richer data on various behavioral margins in order to provide more accurate conclusions on the effect of different types of behavioral changes on the overall elasticity of taxable income.

Bibliography

[1] Aarbu, K. and T. Thoresen. (2001). Income responses to tax changes – Evidence from the Norwegian tax reform. National Tax Journal, 54(2): 319–334.

[2] Blomquist, S. and H. Selin. (2010). Hourly wage rate and taxable labor income responsiveness to changes in marginal tax rates. Journal of Public Economics, 94(11-12): 878–889.

[3] Chetty, R. (2012). Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply. Econometrica, 80(3): 969–1018.

[4] Chetty, R. (2009). Is the taxable income elasticity sufficient to calculate the deadweight loss? The implications of evasion and avoidance. American Economic Journal: Economic Policy, 1(2): 31–52.

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

[6] Feldstein, M. (1999). Tax avoidance and the deadweight loss of the income tax. Review of Economics and Statistics, 81(4): 674–680.

[7] Feldstein, M. (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.

[8] Giertz, S. (2010). The elasticity of taxable income during the 1990s: New estimates and sensitivity analyses. Southern Economic Journal, 77(2): 406–433.

[9] Goolsbee, A. (2000). What happens when you tax the rich? Evidence from executive compensation. Journal of Political Economy, 108(2): 352–378.

[10] Gordon, R. and J. Slemrod. (2000). Are 'real' responses to taxes simply income shifting between corporate and personal tax bases?, in J. Slemrod (ed.), Does Atlas Shrug? The Economic Consequences of Taxing the Rich. New York: Harvard University Press and Russell Sage Foundation, 2000, 240–280.

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

[12] Harju, J. and T. Matikka. (2014). The elasticity of taxable income and income-shifting: What is "real" and what is not? CESifo Working Paper 4905.

[13] Kleven, H. and E. Schultz. (2013). Estimating taxable income responses using Danish tax reforms. Forthcoming, American Economic Journal: Economic Policy.

[14] Kopczuk, W. (2005). Tax bases, tax rates and the elasticity of reported income. Journal of Public Economics, 89(11-12): 2093–2119.

[15] Moffit, R. and M. Wilhelm. (2000). Taxation and the labor supply decisions of the affluent. In J. Slemrod (ed.), Does Atlas Shrug? The Economic Consequences of Taxing the Rich. New York: Harvard University Press and Russell Sage Foundation, 2000, 193–234.

[16] Piketty, T., E. Saez and S. Stantcheva. (2014). Optimal taxation of top labor incomes: A tale of three elasticities. American Economic Journal: Economic Policy, 6(1): 230–271.

[17] Pirttilä, J. and R. Uusitalo. (2005). Veronalennukset vähentävät verotuloja. Kansantaloudellinen aikakauskirja, 4/2005 (in Finnish).

[18] Ramsey, F.P. (1927). A contribution to the theory of taxation. Economic Journal, 37: 47–61.

[19] Riihelä, M., R. Sullström and I. Suoniemi. (2008). Tax progressivity and recent evolution of the Finnish income inequality. Government Institute for Economic Research Discussion Papers 460.

[20] Saez, E. (2010). Do taxpayers bunch at kink points? American Economic Journal: Economic Policy, 2(3): 180–212

[21] Saez, E. (2004). Reported income and marginal tax rates, 1960-2000: Evidence and policy implications. NBER Working Paper 10273.

[22] Saez, E. (2003). The effect of marginal tax rates on income: A panel study of 'bracket creep'. Journal of Public Economics, 87: 1231–1258.

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

[24] Slemrod, J. (1996). High income families and the tax changes of the 1980s: The anatomy of behavioral response. In Feldstein M. and J. Poterba (eds). Empirical Foundations of Household Taxation. Chicago: University of Chicago Press, 1996, 169–192.

[25] Slemrod, J. (1995). Income creation or income shifting? Behavioral responses to the Tax Reform Act of 1986. American Economic Review Papers and Proceedings, 85(2): 175–180.

[26] Slemrod, J. and W. Kopczuk. (2002). The optimal elasticity of taxable income. Journal of Public Economics, 84(1): 91–112.

[27] Thoresen, T. and T. Vattø. (2013). Validation of structural labor supply model by the elasticity of taxable income. Statistics Norway Research Department Discussion Papers No. 738.

[28] Weber, C. (2014). Toward obtaining a consistent estimate of the elasticity of taxable income using difference-in-differences. Forthcoming, Journal of Public Economics.

Appendix

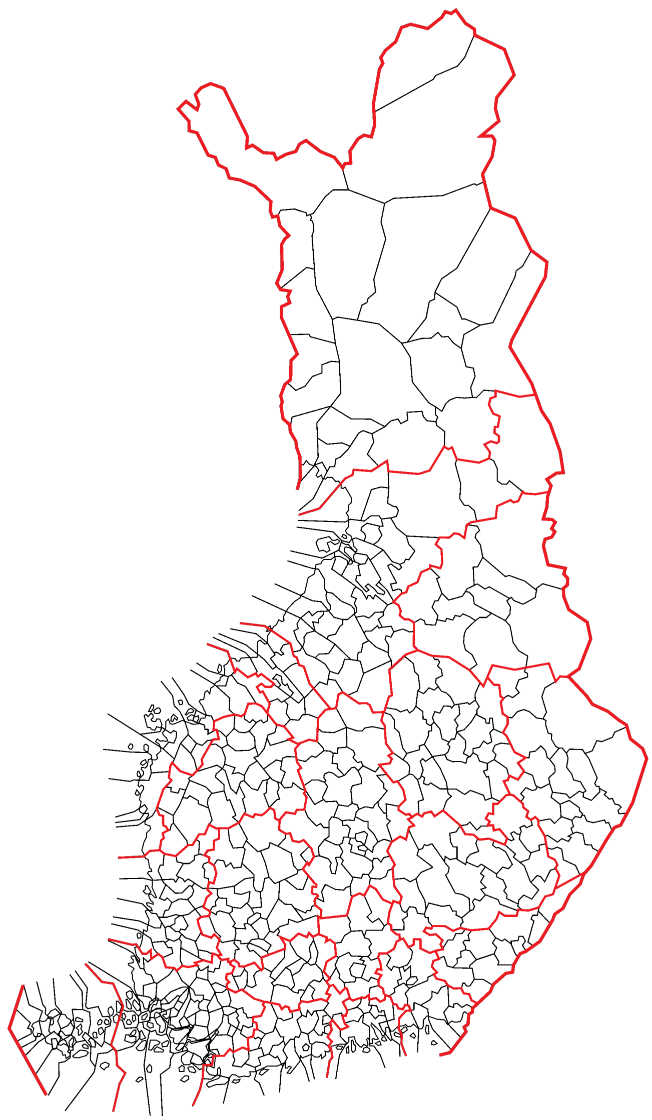


Figure A1. Finnish municipalities and counties in 2007 (Source: National Land Survey of Finland)

Variable	*Description*
Yearly wages (Palkkatulot)	Income regarded as compensation for working in the definition of the Finnish Tax Administration. Includes e.g. wages from primary and secondary jobs and fringe benefits.
Monthly wage rate (Kuukausiansio)	Monthly earnings for full-time workers. Includes wages from regular working hours and overtime work and other additional wage income, including taxable fringe benefits.
Monthly working hours (Kokonaistyöaika)	Working hours for full-time workers. Includes regular working hours (4.35*regular weekly hours) and overtime hours.
Fringe benefits (Luontoisedut)	Taxable fringe benefits for a tax year, including realized options. Fringe benefits include, for example, company cars, phones, apartments, and meals provided by the employer. Usually, the value of taxable fringe benefits is less than the face value of the good, for example in the case of luncheon vouchers provided by the employer. Detailed guidelines are provided by the Finnish Tax Administration.
Work-related expenses (Tulonhankkimisvähennys)	The approved amount of tax-deductible work-related expenses for a tax year. Work-related expenses are deductible from gross earned income. Usual work-related expenses include, for example, the purchasing price of equipment or tools such as computers and professional literature, and office costs if a working space is not provided by the employer. Detailed guidelines are provided by the Finnish Tax Administration.
Commuting expenses (Kodin ja työpaikan välisten matkakustannusten vähennys)	The approved amount of tax-deductible commuting expenses for a tax year. Only the amount exceeding a fixed sum can be deducted (600 euros in 2012). The amount of the deduction is dependent on the mode of transport (public transport, private vehicle or a combination of the two). Detailed guidelines are provided by the Finnish Tax Administration.

Notes: Monthly wage rate and monthly working hours are from the Structure of Earnings statistics (Palkkarakennetilasto), and yearly wages, fringe benefits, work-related expenses and commuting expenses are from the tax statistics produced by the Finnish Tax Administration. The structure of Earnings statistics are based on a sample collected by Statistics Finland. The sample includes both private and public sector workers. The private sector includes only workers in companies with five workers or more. The working hours and wage rate information is mainly based on the situation in October each year. Information on wages and working hours is provided by employers.

Table A1. Subcomponents of taxable income

Year	Taxable income	Tax rate
1995	7,063-9,754	7
	9,754-12,110	17
	12,110-17,155	21
	17,155-26,910	27
	26,910-47,934	33
	47,934-	39
1998	7,737-10,428	6
	10,428-13,119	16
	13,119-18,500	20
	18,500-29,096	26
	29,096-51,466	32
	51,466-	38
2001	11,100-14,296	14
	14,296-19,678	18
	19,678-30,947	24
	30,947-54,661	30
	54,661-	37
2004	11,700-14,500	11
	14,500-20,200	15
	20,200-31,500	21
	31,500-55,800	27
	55,800-	34
2007	12,400-20,400	9
	20,400-33,400	19,5
	33,400-60,800	24
	60,800 -	32

Note: Finnish marks converted to euros before 2002.

TABLE A2. Central government income tax rate schedules in 1995, 1998, 2001, 2004 and 2007

	Mandatory pension insurance contributions	Unemployment insurance contributions	Health insurance payments
1995	4.0%	1.87%	1.9% (3.8% for income > 13,455e)
1998	4.7%	1.4%	1.5% (1.95% for income > 13,455e)
2001	4.5%	0.7%	1.5%
2004	4.6%	0.25%	1.5%
2007	4.3% (5.6% if older than 53 years)	0.58%	1.28%

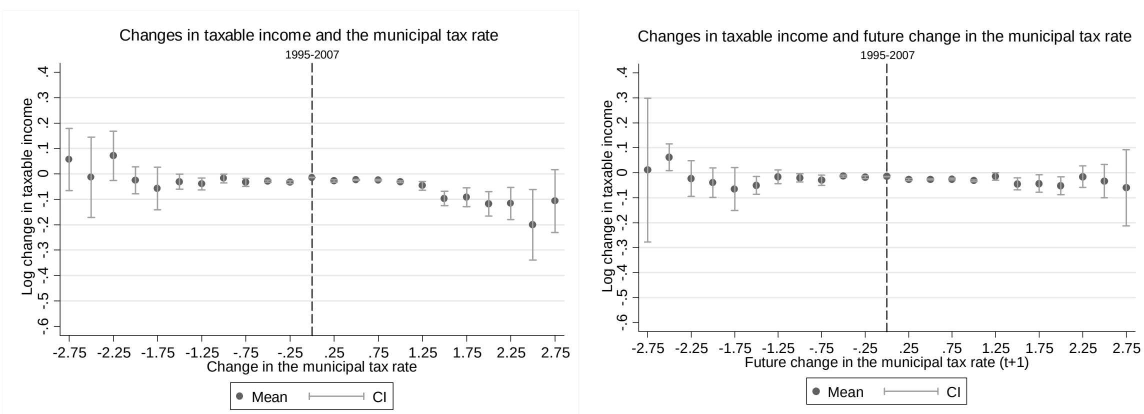
Notes: Pension and unemployment insurance contributions are levied on wage income. Health insurance payments are paid on the basis of taxable income.

TABLE A3. Social security contributions in 1995, 1998, 2001, 2004 and 2007

	Phase-in threshold 1	Phase-in threshold 2	Phase-out threshold	Phase-in rate 1	Phase-in rate 2	Phase-out rate	Max. allowance
1995	3,364	-	13,455	0,05	-	0,05	336
1998	2,523	-	7,232	0,2	-	0,02	925
2001	2,523	-	12,614	0,35	-	0,035	1,648
2004	2,500	7,230	14,000	0,47	0,23	0,04	3,550
2007	2,500	7,230	14,000	0,49	0,26	0,04	3,250

Notes:
Before 2002 (example year 1998): for all wage income earners, a certain amount is deducted from adjusted gross income (gross earned income minus expense deductions) in municipal income taxation as an earned income tax allowance. The allowance is 20% (Phase-in rate 1) of income above 2,523 euros (Phase-in threshold 1). The maximum amount of the allowance is 1,648 euros (Max allowance). For income above 7,232 euros (Phase-out threshold), the allowance decreases by 2% (Phase-out rate). *After 2002 (example year 2004):* the allowance rate is 47% (Phase-in rate 1) of income between 2,500 euros (Phase-in threshold 1) and 7,230 euros (Phase-in threshold 2). The allowance rate is 23% (Phase-in rate 2) of income above 7,230 euros. The maximum amount of the allowance is 3,550 euros (Max allowance). For income above 14,000 euros (Phase-out threshold), the allowance decreases by 4% (Phase-out rate).

TABLE A4. Earned income tax allowance in municipal taxation in 1995, 1998, 2001, 2004 and 2007



Notes: The baseline sample includes observations where base-year taxable income is above 20,000 euros. Pensioners and people under the age of 24 and over the age of 60 are not included in the sample. Also, the sample is limited to individuals whose absolute change in log taxable income between t and $t+1$ is below 8.5, and whose marital status is unchanged between the two years.

FIGURE A2. Log changes in taxable income and changes in municipal tax rates, movers included

Individuals					
Variable	Obs.	Mean	Std. Dev.	Min	Max
Taxable earned income	3,116,040	20,893	28,275.61	0	1.88e+07
Gross earned income	3,116,040	24,726	29,134.99	0	1.89e+07
Total taxable income (earned+capital income)	3,116,040	22,080	48,116.18	0	3.35e+07
Wage income	3,123,447	20,287	29,522.76	0	1.88e+07
Commuting expense deduction	3,111,906	332.60	859.57	0	7,000
Work-related expense deduction	3,111,906	107.35	699.12	0	268,400
Fringe benefits	3,123,447	447.50	19,635.25	0	1.87e+07
Monthly wage	1,398,846	2,320	1,233.83	0	221,100
Monthly working hours	1,397,291	156	32.77	0	250
Age	3,127,819	42.06	9.46	25	60
Female	3,127,819	0.50	0.5	0	1
Size of the household	3,105,782	3.56	1.65	1	25
Municipal tax rate	3,127,340	17.95	0.77	16	21
Marginal tax rate	3,127,340	0.393	0.13	0	0.668
Municipalities					
Variable	Obs.	Mean	Std. Dev.	Min	Max
Municipal income tax rate	5,733	18.29	.77	16	21
Average individual taxable income	5,734	12,933.40	2,868.85	5,509	56,055
Net loans position (per capita)	4,690	1,925.73	2,102.40	-1	30,453
Employment rate	5,734	0.617	.075	0.391	0.840
Unemployment rate	5,726	0.141	.06	0.004	0.400
Net migration	5,676	-0.003	.011	-0.074	0.061

Notes: Income variables in 2007 euros. Individual statistics are calculated for individuals aged 25-60. Individual maximum values shown in the table are rounded to the nearest 10 or 100. The average taxable income within a municipality includes all individuals.

Table A5. Summary statistics (whole data), 1995-2007

	(1) Movers included	(2) Gross earned income	(3) Income-weighted	(4) Men	(5) Women	(6) Low income (10k-25k)
Elasticity	0.629***	0.198*	0.274**	0.234	0.431**	0.333**
	(0.154)	(0.119)	(0.135)	(0.158)	(0.209)	(0.140)
Observations	426,714	413,480	413,482	268,805	144,677	501,725
	(7) Mid. income (25k-40k)	(8) High income (>40k)	(9) Without capital city area	(10) 1-year	(11) 5-year	(12) 1996-2001 and 2002-2007
Elasticity	0.128	0.973**	0.248	0.431***	-0.104	0.670**
	(0.179)	(0.492)	(0.174)	(0.084)	(0.247)	(0.266)
Observations	187,834	72,874	297,919	932,319	158,609	108,023
	(13) Gruber and Saez instrument	(14) OLS	(15) First-stage	(16) Reduced-form		
Elasticity	-0.234***	-2.878***	1.425***	0.379**		
	(0.027)	(0.019)	(0.046)	(0.180)		
Observations	413,122	654,438	413,482	413,482		

Robust municipal-level clustered standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

All regressions include year, county and county-year fixed effects, individual base-year controls and municipal controls. Column (13) also includes income splines.

Notes: The dependent variable is log change in taxable income if not mentioned otherwise. The dependent variable in (2) is gross earned income subject to taxation. The capital city area in column (9) includes Helsinki, Espoo, Vantaa and Kauniainen. Column (12) shows the results for the pooled regression of two six-year differences (1996-2001 and 2002-2007). The estimation sample in column (13) follows the baseline sample where individuals with municipal tax rate changes in the middle of the difference are not included. The dependent variable in (15) is the change in the overall net-of-tax rate. The F-test statistic in (15) is 824.45 (0.000).

Table A6. Alternative specifications

CHAPTER 5

Unwilling, Unable or Unaware? The Role of Different Behavioral Factors in Responding to Tax Incentives[1]

Abstract. This paper characterizes how various behavioral factors affect individual responses to different tax incentives. This is important because different reasons for responding and not responding might have different policy implications and welfare conclusions. Our analysis compares the empirical significance of the inability to respond to tax incentives and unawareness of tax rules. Using population-wide Finnish panel data, we estimate behavioral responses to kinks and notches within the tax and transfer system among similar or even the same individuals. We find that taxpayers do not respond at all to small incentives induced by kink points, but do respond to larger incentives induced by notches. The patterns of responding suggest that some taxpayers are unable to respond even to large incentives, and that unawareness might also affect the negligible response to kink points.

Keywords: income taxation, income transfers, behavioral responses, frictions
JEL codes: H21; H24; H31

5.1. Introduction

Existing studies find varying responses to similar income tax incentives (see Saez et al. 2012 and Meghir and Phillips 2010 for surveys). While these divergent responses are traditionally explained by heterogeneous preferences, recent literature adds optimization frictions to explain differences in observed behavior. Optimization frictions potentially prevent taxpayers from fully responding to tax incentives according to their underlying preferences (Chetty 2012, Kleven and Waseem 2013). One example of these frictions is job search costs (Chetty et al. 2011), which reduce the willingness to switch jobs when tax incentives change. Other examples in previous literature are insufficient knowledge of tax rules, inattention and the salience of tax regulations (Chetty

[1]This essay is joint work with Tuomas Kosonen (Government Institute for Economic Research VATT).

et al. 2013, Chetty and Saez 2013, Chetty et al. 2009). Although different institutional settings could feature different optimization frictions, the literature has not thus far systematically addressed their role in explaining heterogeneous responses to tax incentives.

Understanding the role of various optimization frictions has important policy implications. Different frictions might imply different patterns of responding to similar tax incentives (Reck 2014, Chetty et al. 2009 and 2007). For example, when observed behavioral responses are attenuated by the inability to respond immediately because of rigid labor demand, we would expect individuals to adjust their behavior in the future, and this adjustment might cause notable welfare losses. On the other hand, when responses are attenuated by unawareness of tax regulations or inattention, it is not clear whether individuals would be more aware or attentive over time (and change their behavior accordingly). Therefore, if individuals do not even know or understand that there has been a change in tax incentives, it is not certain that welfare losses occur either. The extent of the welfare loss matters, among other policy-relevant issues, for the design of optimal tax schedules.

In this paper we study to what extent and in what manner taxpayers respond to different tax incentives. We use local variation in tax incentives created by different tax and transfer schemes. First, we study discontinuous jumps in marginal income tax rates (kinks). Under standard labor-leisure preferences, if individuals bunch at these kink points, it can be seen as clear evidence that marginal tax rates affect the behavior of taxpayers (Saez 2010). However, if taxpayers do not bunch, it remains an open question whether the tax incentives are inherently not large enough to induce responses, optimization frictions eliminate the observed response, or the underlying structural model is not correctly specified.

Second, we utilize a stronger variation in tax incentives created by a means-tested income transfer, the study subsidy. In Finland, all university students can apply for a substantial study subsidy (approx. 500 euros/month). However, earning income above an income limit results in losing part of this subsidy, which creates a jump in the average tax rate, called a notch (see Slemrod 2010).

Similarly, bunching at a notch in the tax schedule provides clear evidence that individuals respond to tax incentives. However, there are two key differences between responses to notches and kinks. First, notches create much stronger variation in incentives than kinks. Thus comparing responses to kinks with responses to notches allows us to outline the role of the strength of the tax incentive in explaining taxpayer responses. Second, according to standard economic theory, taxpayers should never locate themselves just above the notch where they *lose* disposable income compared to the notch point. Utilizing this so-called dominated region and the shape of the income distribution around the notch allow us to characterize the role of optimization frictions. There is no such dominated region of behavior associated with kink points.

We use register-based panel data on all Finnish taxpayers from 1999-2011. The data include detailed tax and transfer variables from official registers. These data allow us to accurately analyze bunching behavior associated with various kinks and notches.

One particular advantage is that we can compare how similar or even the same taxpayers react to different tax incentives, whilst keeping other institutional features constant. Also, the large and extensive data set allows us to conduct the bunching analysis for various subgroups, such as wage earners, self-employed individuals and students with different characteristics.

Our findings support the view that frictions play an important role in explaining taxpayer responses to tax incentives. First, we do not find any bunching at kink points. This result holds for any tax rate kink and for any subgroup. This implies that either the structural elasticity is small or optimization frictions play an important role, or both.

Interestingly, we find no bunching at kink points for the self-employed (sole proprietors and partners of partnership firms). However, we find that self-employed individuals bunch actively at round numbers (e.g. multiples of 10,000 euros) of personal gross earned income. Thus they are at least somewhat able to alter their reported income, but despite this they choose not to report income that is close to tax rate kink points. In general, the ability to affect reported income suggests that inability to respond does not prevent the self-employed from bunching more prominently at the marginal tax rate kink points.

Second, we find that income notches related to the study subsidy system create significant bunching behavior among students. This indicates that given strong enough incentives, taxpayers do react to local variation in the tax schedule. However, although the bunching behavior is evident, the local changes in tax incentives are so large that the implied observed earnings elasticities are in general small (<0.1).

As for other groups, we do not find any bunching at the marginal tax rate kink points for students. However, bunching at the notch reveals that at least some students are able to respond to tax incentives, and thus no bunching at kink points is not completely driven by the inability to respond to any local tax incentive.

Furthermore, despite the distinctive bunching behavior, we find that many students are located in the dominated region just above the notch. This is compelling evidence in favor of notable optimization frictions (Kleven and Waseem 2013). To characterize the source of the friction, we turn to institutional features associated with the study subsidy and the labor market students are in. We hypothesize that most of the friction is due to the inability to respond. Compared to the self-employed, it is difficult for students to choose or report any income they want. Instead, they have a limited number of wage and hours contracts to choose from, and it could be costly to search for a new job or to stop working abruptly at a certain point of time during the year when the income limit is reached.

We hypothesize that unawareness of study subsidy rules is not the main friction. Notches created by the study subsidy system are fairly simple and transparent. First, students need to apply for the subsidy, which makes it an active choice. Second, when they get their acceptance decision, the income limits are stated in the notification letter. Third, the Social Insurance Institution reclaims the subsidy if students earn income above the limit. In comparison, taxpayers might not be aware of marginal tax rates, or correctly understand what changes in marginal tax rates indicate. The fact that we do not find any bunching at kink points for students or even for the self-employed supports this view.

In order to further support the view that inability to respond is the source of frictions for students as opposed to unawareness of the rules, we study how changes in the location of the notch point affect the behavior of students The income limits were not changed even in nominal terms for several years. However, in 2008, the income limits were adjusted upwards by 30%. We find that students start bunching at the new income limit immediately, indicating that they are aware of the rules.

In addition to optimization frictions, this study contributes to the literature on observed responses to kinks and notches. This study is the first to analyze bunching at marginal tax rate kink points in Finland. Many previous studies find no or only little bunching at the kink points of the marginal tax rate schedule for wage earners, but significant and sharp bunching for the self-employed (Saez 2010, Bastani and Selin 2014 and Chetty et al. 2011). One intriguing finding in this study compared to the earlier literature is that we find only negligible if any bunching at marginal tax rate kink points for the self-employed.

Kleven and Waseem (2013) show that wage earners bunch actively at income tax notches in Pakistan. We add to this study by estimating responses to income notches in a developed country where the tax system is strongly enforced, and thus the responses are more related to labor supply decisions as opposed to reporting behavior. Other existing evidence on responses to notches comes from a range of institutions, for example the medicaid notch in the US (Yelowitz 1995), eligibility for in-work benefits in the UK (Blundell and Hoynes 2004, Blundell and Shepard 2012), social security and financial incentives in retirement rules (Gruber and Wise 1998, Manoli and Weber 2011), and car taxes affecting the fuel economy of cars (Sallee and Slemrod 2012).

This paper proceeds by presenting the relevant institutions in Section 5.2. In Section 5.3 we present the conceptual background to responding to kinks and notches, and discuss the role of behavioral frictions. We then present the empirical methodology and data in Section 5.4. Section 5.5 presents the results. Section 5.6 discusses the implications and concludes the study.

5.2. Institutions

5.2.1. Income taxation and marginal tax rate kink points. We study the marginal tax rate (MTR) kink points created by the central government income tax schedule.[2] Small amounts of earned income are not taxed by central government. The first kink appears at a point where the central government tax rate first applies. After the central government tax rate is first applied, it increases in a stepwise manner. This results in 4-6 kink points in the MTR schedule, depending on the year in question.

Various kink points are associated with MTR increases between 4-11 percentage points. At the first income threshold, there is a clear increase in the overall MTR.

[2]The Finnish income tax system comprises three components: progressive central government income taxes, proportional municipal taxes and mandatory social security contributions. The average municipal income tax rate is 18.3%, and the average social security contribution rate is 5.1% (in 1999-2011). In general, municipal income taxation and social security contributions do not induce kink points since they are proportional. The main exception is the municipal earned income tax allowance, which will be briefly discussed in Section 5.5. Since 1993, Finland has applied a dual income tax system. In dual income taxation, earned income (wages, fringe benefits, pensions etc.) is taxed at a progressive tax schedule, and capital income (interest income, dividends from listed corporations etc.) is taxed at a flat tax rate. In this study we focus on the details of the earned income tax schedule. However, the dual income tax system affects the tax rules of self-employed individuals, which we discuss at the end of this section.

In 1999-2011, the increase in the MTR associated with the first income threshold has varied between 6-14 percentage points, which relates to a 22-53% decrease in the overall net-of-tax rate (1-MTR) on average (excluding employer social security payments). In addition to the first kink point, the last kink involves the most salient and distinctive increase in the MTR. The last kink point is associated with a 6-9 percentage point increase in the MTR, and a 9-16% decrease in the overall net-of-tax rate.

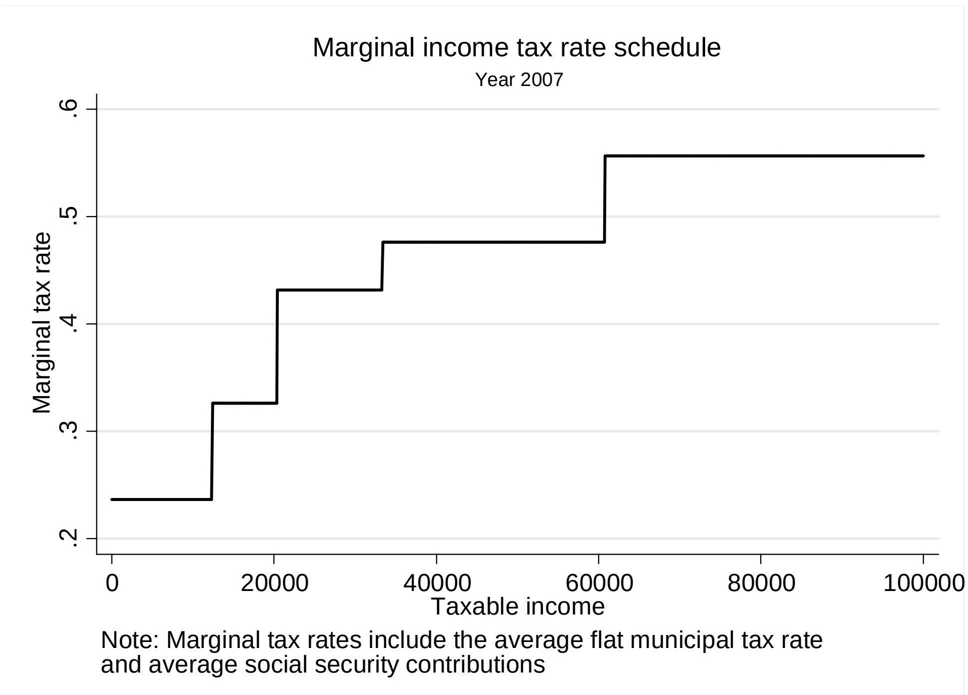


FIGURE 1. Marginal income tax rate schedule (year 2007)

As an example, Figure 1 presents the marginal income tax rate schedule for 2007. The Figure illustrates the discontinuous changes in the income tax rate at different levels of taxable income. Taxable income is the basis of central government taxation, and it is roughly defined as gross earned income minus deductions.

In order to take into account the general increase in wages and other prices over time, the nominal income thresholds have moved upwards over time. However, increases in the income thresholds are not tied to any price or wage index, and are announced by the government on annual basis. Table A1 in the Appendix presents the nominal MTR schedules of central government income taxation in 1999-2011. Figure A1 in the Appendix presents the overall nominal average marginal income tax rates in 1999, 2007 and 2011 (including average municipal tax rates and social security contributions).

In addition to wage earners, we study the behavior of self-employed individuals. In this study self-employed individuals include sole proprietors and partners of partnership firms (all non-corporate entrepreneurs in Finland). The annual reported income of these individuals is based on the reported profits (earnings-costs) of their firms. In the Finnish dual income tax system, with separate tax rate schedules for earned and capital income, these profits are mechanically divided into capital income and earned income components.

Profits are divided into capital income and earned income according to the net assets of the firm (assets-liabilities from the year before). An amount corresponding to 20% of net assets is considered flat-taxed capital income, and any profits exceeding this amount are progressively taxed as earned income.[3] In the case of zero or negative net assets, the profits are taxed completely as earned income.

As an example, consider a self-employed individual who is the sole owner of a firm, and has net assets of 100,000€ and profits of 30,000€. In this case, 20,000€ of the profits are flat-taxed, and the remaining 10,000€ are taxed on a progressive tax rate schedule, illustrated in Figure 1. Without any net assets, the whole 30,000€ is taxed as earned income.

Intuitively, all self-employed individuals face similar *local* incentives within the earned income MTR schedule as regular wage earners. Even though profits are partly flat-taxed, the kink points of the earned income tax schedule provide similar marginal

[3]The flat capital income tax rate was 28% in 1999, and 29% in 2000-2004. In 2005-2011, the rate was reduced to 28%.

changes in incentives. Furthermore, as the maximum amount of flat-taxed capital income is predetermined based on net assets from the year before, there is no possibility for static income-shifting between tax bases among sole proprietors and partners of partnership firms in Finland.

5.2.2. Study subsidy. In Finland, all students enrolled in a university or polytechnic can apply for a monthly-based study subsidy.[4] The maximum amount of the subsidy is 461€ per month in the academic year 2006/2007.[5] Students can apply for the subsidy for a limited number of months per degree (max. 55 months).

The study subsidy is typically available when a student is accepted to study for a university or college degree. The default number of study subsidy months per study year is 9 (fall + spring semester), which most students also receive. Study subsidy eligibility depends on academic progress[6], and is reduced if the yearly gross earned income of the student is too high.

Students can earn a certain amount of gross income (earned income + capital income) per calendar year without an effect on the study subsidy. With the typical 9 months of the subsidy per calendar year, the annual gross income limit is 9,260€ (in 2006/2007). Students can alter the number of subsidy months from the default 9 months by making an application beforehand, or by returning already granted subsidies

[4] The study subsidy is intended to enhance equal opportunities to acquire higher education, and to provide income support for students who often have low disposable income. In Finland, university education is publicly provided, and consequently there are no tuition fees. A large proportion of individuals receive higher education in Finland (ca. 40% of individuals aged 25-34 have a degree), and the study subsidy program is widely used among students.

[5] The full study subsidy includes a study grant and housing benefit. The standard study grant is 259€/month (in 2006/2007). The housing benefit depends on rent payments and other housing details. The maximum housing benefit is 202€/month (in 2006/2007). In addition to the study subsidy, students can apply for repayable student loans which are secured by the central government.

[6] The academic progress criteria requires that a student completes a certain number of credit points per academic year in order to be eligible for the subsidy.

by the end of March in the next calendar year. More study subsidy months decreases the income limit, and less study subsidy months increases it.[7]

The gross income limit in the study subsidy program creates a significant notch in the tax system. If the income limit is exceeded, one month‘s study subsidy is reclaimed by the Social Insurance Institution. An additional month of subsidy is reclaimed for an additional 1,010€ of gross income over the threshold.

Students face large local incentives not to exceed the income limit. Since earning just a little over the limit results in losing one study subsidy month, this results in an implied marginal tax rate of over 100% just above the notch. Thus the study subsidy notch induces a strictly dominated region above the notch where students can earn more disposable income by decreasing their gross income level.

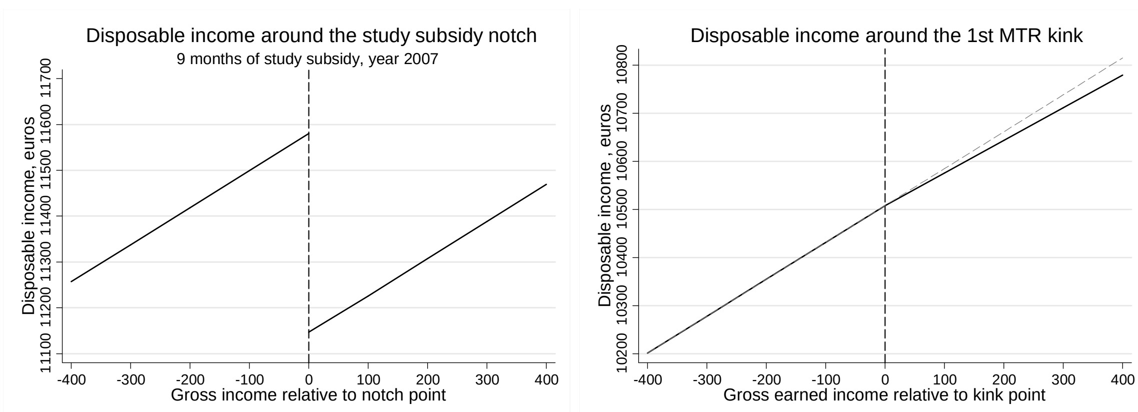


FIGURE 2. Disposable income around the study subsidy notch (left-hand side) and the first MTR kink point (right-hand side), year 2007

The left-hand side of Figure 2 illustrates the effect of the study subsidy notch on disposable income with the standard case of 9 study subsidy months (in 2007). In the Figure, the vertical axis denotes disposable income including the subsidy, and the horizontal axis denotes gross income relative to the notch point (9,260€). The Figure

[7]The formula for the annual gross income limit is the following: 505€ per study subsidy month plus a fixed amount of 170€, and 1,515€ per month without the study subsidy (in 2006/2007).

shows that once the gross income limit is exceeded, reclaiming of the study subsidy causes a dip in disposable income. At the margin, earning 100€ above the threshold results in a loss of 360€ in disposable income.

The right-hand side of Figure 2 illustrates the effect of the first marginal income tax rate kink point on disposable income. Earning income after the kink point results in less disposable income than before the kink. For example, 100€ of gross income above the kink results in 9€ less disposable income than below the kink.

Figure 2 highlights that the difference between the study subsidy notch and the MTR kink points is notable. Even though kink points also change the incentives at the margin, the effect of the study subsidy notch is significantly larger.

The study subsidy program was reformed in 2008. The main outcome of the reform was that the income limits were increased by approximately 30%. The default income limit for 9 study subsidy months increased from 9,260€ to 12,070€[8] In addition, the monthly study subsidy was increased from 461€ to 500€ per month. In general, other details of the system were not changed, including the academic criteria and the loss of one month's subsidy if the income limit is exceeded.[9] Finally, Table A2 in the Appendix shows the income limits for differing number of study subsidy months, and the relative loss incurred when the income limit is exceeded both before and after the reform.

5.3. Conceptual framework

5.3.1. Behavioral responses to kinks and notches. We analyze taxpayer responses to kinks and notches in the tax schedule using a static model that closely follows Saez (2010) and Kleven and Waseem (2013). In short, the model shows that if behavioral responses are notable, we should find individuals bunching in the income

[8]After 2008, the gross income limits are 660€ (before 505€) per study subsidy month plus a fixed amount of 220€ (170€), and 1,970€ (1,515€) per month when no study subsidies are claimed.
[9]After 2008, an additional month of the subsidy is reclaimed for an additional 1,310€ of gross income over the threshold, compared to 1,010€ before 2008.

distribution at the kink and notch points. We first analyze behavioral responses without frictions and then discuss how different frictions alter the baseline bunching formula.

We assume that individuals have a quasi-linear utility (no income effects). Individuals have homogenous tastes and labor supply elasticities but different abilities, which gives rise to the shape of the income distribution. The iso-elastic utility function is of the form

$$u(c,z) \quad = \quad c - \frac{1}{1+1/e}\frac{z}{n}^{1+1/e}$$

where c is consumption, z is gross earnings, e is the earnings elasticity and n is ability.

Individuals maximize utility with respect to a budget constraint $c = z(1-t) + R$, where R denotes virtual income. We focus on linear income tax rates t to simplify the problem. Maximizing utility with respect to the constraint gives the following earnings supply function

$$z = n(1-t)^e$$

We assume that there is a continuous distribution of abilities, giving rise to a density function $f(n)$ and a distribution function $F(n)$. For a baseline tax system which is linear and has no kink points, there is an earnings distribution associated with a density and distribution function, $H_0(z) = F(z/(1-t)^e)$ and $h_0(z) = H_0'(z) = f(z/(1-t)^e)/(1-t)^e$.

Next, we look at how kinks and notches transform the underlying distributions with no kinks or notches. For kink points, consider a small increase in the marginal tax rate, dt, at a point $z = k$. At k income is taxed at a tax rate t_1, and above the kink point the tax rate is $t_2 = t_1 + dt$. Individuals who were previously located at the kink do not need to change their behavior, but individuals above the kink face a higher tax rate than before. dz denotes the behavioral changes in gross earnings as a response to the increased tax rate. In terms of the earnings elasticity e, the behavioral responses can be written as

$$\frac{dz}{k} = e\frac{dt}{1-t_1}$$

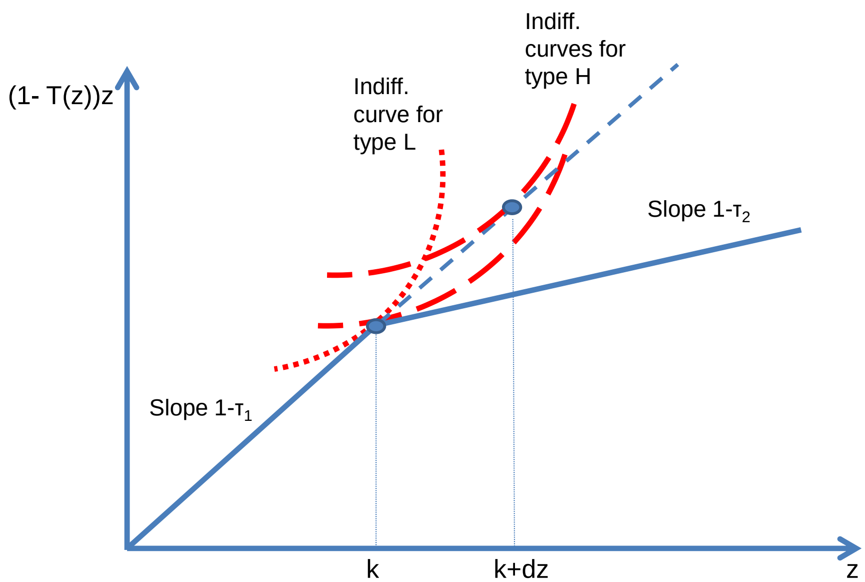


FIGURE 3. Bunching at a kink point

Figure 3 illustrates the bunching effect in the absence of frictions. The vertical axis denotes net-of-tax income, and the horizontal axis denotes pre-tax income. The straight blue lines illustrate the tax rates, and the curvy red lines the indifference curves. As a result of the behavioral response to the introduced kink point k, individuals located within the income interval $(k, k + dz)$ now bunch at k. In the Figure, an individual of type H is the highest pre-tax income individual to move to the kink point. Individuals further up in the income distribution $z > k + dz$ do not move to the kink point, and individuals originally located below or at the kink point (Type L) do not change their behavior either. Thus we can express the extent of bunching behavior as $B(dz) = \int_k^{k+dz} h_0(z)dz$.

Notches can be analyzed in a similar fashion. The tax schedule above the notch point at $z = j$ is characterized as $t + \Delta t$. For $z <= j$, the income tax rate is t. When $z > j$, income is taxed at a tax rate t plus an additional tax of Δt. In the case of income transfers with income limits, Δt can be thought of as the forfeit transfer when the income limit is exceeded.

Notches create a so-called dominated region just above the notch point where individuals can increase their net-of-tax income by moving to the notch point and earning *less* pre-tax income. Under normal preferences and in the absence of any frictions, no individuals should locate themselves within the dominated region.

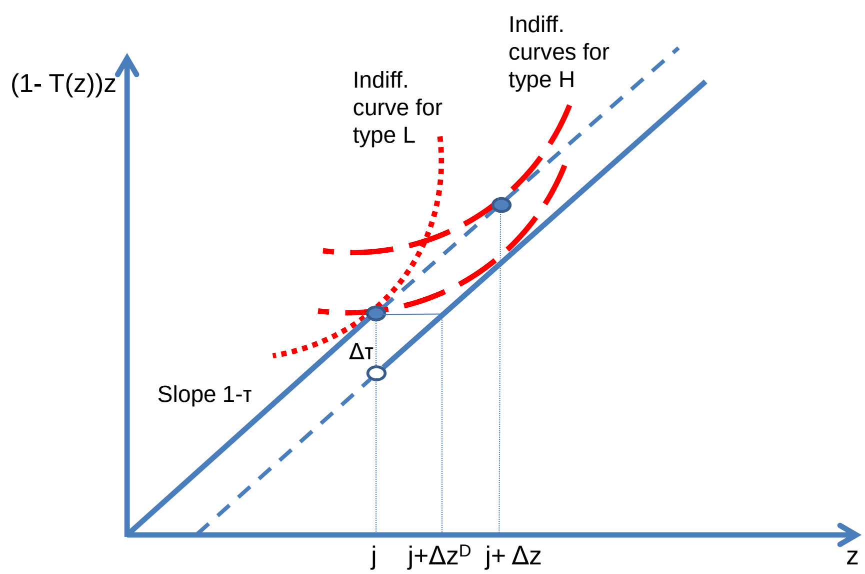


FIGURE 4. Bunching at a notch point

Figure 4 illustrates the bunching effect related to notches. Individuals located within $(j, j+\Delta z)$ will bunch at the notch point, and type H individual is the last to move to the notch. Thus type H individual represents the marginal buncher with the highest pre-tax income before the implementation of the notch. The bunching behavior is denoted as $B(\Delta z) = \int_j^{j+\Delta z} h_0(z)dz$.

In the figure, the dominated region is denoted as $(j, j+\Delta z^D]$. Throughout the paper, we define the dominated region such that the upper limit of the region is a point where the net-of-tax income equals the net-of-tax income at the notch. By definition, all points between the notch and the upper limit of the dominated region produce less net-of-tax income compared to the notch point.

5.3.2. Earnings elasticities based on observed bunching. Following Saez (2010), using the expression for excess bunching $B(dz)$ along with the taxable income elasticity formula by Feldstein (1999), we can express the local average elasticity of taxable income (ETI) at the kink point in proportion to the number of individuals bunching at the kink point

$$e(k) \simeq \frac{B(dz)}{k \times h_0(k) \times log(\frac{1-\tau_1}{1-\tau_2})} \tag{5.3.1}$$

In equation (5.3.1), k is the kink point, $h_0(k)$ denotes the counterfactual density in the absence of the kink point, and $(1-\tau_1)$ and $(1-\tau_2)$ denote the net-of-tax rates below and above the kink point, respectively.

Intuitively, larger $B(dz)$ indicates larger behavioral responses and larger local elasticity, and vice versa. Also, with a given $B(dz)$ and $h_0(k)$, a smaller difference between the tax rates τ_1 and τ_2 indicates larger local elasticity. As underlined in Feldstein (1999), this elasticity measure is directly proportional to the excess burden of the income tax. Thus, in the absence of frictions, we can measure the excess burden with $e(k)$.

As the behavioral response to a notch is related to changes in average tax rates rather than marginal tax rates, deriving the implied elasticity using excess bunching at notches is less straightforward. However, the earnings elasticity at a notch can be approximated in terms of the excess mass at the notch point and the implied change in the marginal tax rate above the notch.

We approximate the earnings elasticity at the study subsidy notch using a similar approach as Kleven and Waseem (2013). We derive an upper-bound reduced-form earnings elasticity by relating the earnings response of a marginal buncher at $j + \Delta z$ to the implicit change in tax liability between the notch point j and $j + \Delta z$. The marginal buncher represents the individual with the highest income to move to the notch point, compared to a counterfactual state in the absence of the notch (see Figure 4).

Intuitively, this approach treats the notch as a hypothetical kink which creates a jump in the implied marginal tax rate. More formally, the reduced-form earnings elasticity is calculated with a quadratic formula

$$e(j) \approx (\triangle z/j)^2/(\triangle t/(1-t)) \tag{5.3.2}$$

where $(1 - t)$ is the net-of-tax rate at the notch, and $\triangle t$ defines the change in the implied marginal tax rate for the marginal buncher with an earnings response of Δz.

5.3.3. Frictions. In the forthcoming analysis, we decompose behavioral frictions into two broadly defined components: unawareness of tax rules and regulations and inability to respond to tax incentives. Unawareness of tax rules covers the lack of knowledge that taxpayers might have about tax regulations. This includes both pure inattention to tax rules and failure to understand them even when general knowledge about tax regulations is available. For example, taxpayers might not know that kink or notch points even exist, or not know the correct income base determining their location in the income tax schedule.

Unawareness also includes any mistakes that taxpayers might make in interpreting the actual incentives. A well-known example is the confusion between marginal and average tax rates (see e.g. Chetty and Saez 2013, and Liebman and Zeckhauser 2004). Misunderstanding of marginal changes in incentives might induce individuals not to respond to local changes in incentives.

The inability to respond covers a range of reasons why taxpayers are not able to flexibly respond to tax incentives. These include factors constraining behavioral responses even when taxpayers are aware of local incentives. The inability to respond might stem from institutional factors as well as individual constraints. For example, due to fixed long-term contracts, wage earners might not be able to alter their working hours easily. Also, it might be costly to search for a new job providing more suitable working hours and wage rates in terms of tax incentives (see e.g. Chetty et al. 2011). Intuitively, when inability frictions are present, large local changes in incentives should produce more observed bunching than small changes, since it is on average more profitable to overcome the inability friction when payoffs from changing behavior are larger (Chetty 2012).

In general, compared to underlying structural responses in the frictionless benchmark (Figures 3 and 4), frictions attenuate the observed behavioral responses. However, different frictions potentially cause different patterns of observed behavior. If individuals are both aware of tax changes and able to respond to them, we should observe sharp bunching at kinks and notches if the underlying elasticity is significant. If some or all individuals are unaware of tax rules, this would either mitigate or eliminate the sharp response. In contrast, if individuals are aware but not able to fully respond, the observed bunching response would not be sharp but more scattered around kinks and notches.

Furthermore, different frictions imply different reasons for responding or not responding to tax incentives. Consequently, different frictions hold potentially different policy implications and long-run welfare conclusions. We discuss these in more detail when we interpret and discuss the results in Section 5.6.

Finally, we include optimization frictions in the theoretical analysis. Since all frictions have an *a priori* similar effect on average responses in a cross-sectional context, we denote frictions by a single term a, $0 < a < 1$. The higher a is, the larger the frictions are and the less individuals respond to tax incentives. The fraction of individuals responding to tax incentives in the presence of frictions is denoted as $(1 - a)$.

After including the frictions, the two bunching formulas become $B_a(dz) = \int_k^{k+dz}(1-a)h_0(z)dz$, and $B_a(\Delta z) = \int_j^{j+\Delta z}(1 - a)h_0(z)dz$. It is evident that bunching behavior is reduced when frictions exist. Nevertheless, the behavioral response in the absence of frictions might still be non-negligible, giving rise to a baseline long-run structural earnings elasticity.

5.4. Empirical methodology and data

5.4.1. Bunching at kinks and notches. With both kinks and notches it is straightforward to verify visually whether there is bunching or not. The challenge is in estimating the size of the excess mass in relation to the counterfactual state of no kinks or notches. In short, the excess mass of individuals at a kink or a notch is estimated by comparing the actual density function around the discontinuity point k to a smooth counterfactual density. The counterfactual density function describes what the income distribution at the notch or kink would have looked like without a change in the tax rate. The bunching method implicitly assumes that individuals in the neighborhood of a kink or a notch are otherwise similar except that they face a different slope or shape in their budget set.

Due to imperfect control and uncertainty about the exact amount of income in each year, the usual approach is to use a "bunching window" around k to estimate the excess mass (see Saez 2010). Thus when analyzing kink points, we compare the density of taxpayers within an income interval $(k-\delta_L, k+\delta_H)$ to an estimated counterfactual density within the same income range. δ_L denotes the lower income limit on the left of the kink, and δ_H denotes the income limit above k.

We follow Chetty et al. (2011) to estimate excess bunching at kink points. The counterfactual density is estimated by fitting a flexible polynomial function to the observed density function, excluding the region $[\delta_L, \delta_H]$ from the regression. First, we re-center income in terms of the discontinuity point, and group individuals into small income bins of 100€. Next, we estimate a counterfactual density by regressing the following equation

$$c_j = \sum_{i=0}^{p} \beta_i (z_j)^i + \sum_{i=\delta_L}^{\delta_H} \eta_i \cdot \mathbf{1}(z_j = i) + \varepsilon_j \tag{5.4.1}$$

and by omitting the bunching window $(k-\delta_L, k+\delta_H)$ from the regression. In equation (5.4.1), c_j is the count of individuals in bin j, and z_j denotes the income level in bin j. The order of the polynomial is denoted by p. Thus the fitted values for the counterfactual density are given by

$$\hat{c}_j = \sum_{i=0}^{p} \beta_i (z_j)^i \tag{5.4.2}$$

The relative difference between the observed individuals and the counterfactual density within the bunching window defines the excess bunching. More formally, for kink points, excess bunching is calculated as

$$\hat{b}(k) = \frac{\sum_{i=\delta_L}^{\delta_H} (c_j - \hat{c}_j)}{\sum_{i=\delta_L}^{\delta_H} \hat{c}_j / (\delta_H - \delta_L + 1)} \tag{5.4.3}$$

As in the earlier literature, parameters δ_L, δ_H and p are determined visually and based on the fit of the model. In general, our results are not very sensitive to the choice of the omitted region or the degree of the polynomial.[10]

The method for analyzing excess mass at notches is based on similar principles. The main difference with notches is that the excess mass should locate below the notch and not as a diffuse mass around both sides of it. Thus in the case of notches, the excess bunching is measured by comparing the observed distribution and the counterfactual within the interval $(k - \delta_L, k)$, where δ_L is the lower limit of the interval and k refers to the notch point.

With notches it is less straightforward to define the income limit above the notch point when estimating the counterfactual density. We follow Kleven and Waseem (2013) and define the upper limit for the excluded region δ_H such that the excess mass $\hat{b}_E(k) = (\sum_{i=\delta_L}^{k} c_j - \hat{c}_j)$ equals the missing mass above the notch $\hat{b}_M(k) = (\sum_{z>k}^{\delta_H} \hat{c}_j - c_j)$. This procedure is implemented by starting from a small value of δ_H and increasing it incrementally until $\hat{b}_E(k) \approx \hat{b}_M(k)$. Intuitively, this convergence condition implies that the excess mass below the notch comes from the missing mass above the notch, and that we can define the earnings response $\triangle z$ and the marginal buncher using the estimated excess mass. This definition for δ_H also denotes the upper bound of the excluded range (Kleven and Waseem 2013).

In order to assess frictions related to responding to notches, we measure the relative proportion of individuals who locate at the dominated region just above the notch. Following Kleven and Waseem (2013), individuals at the dominated region are inherently not able to respond to the notch because of frictions, as these individuals would have

[10]Chetty et al. (2011) adjust the counterfactual density above the kink such that it includes the excess bunch at the kink, making the estimated counterfactual density equal to the observed density. Due to the small observed excess bunching at kink points, this has only a trivial effect for our empirical analysis, and thus we estimate the counterfactual for kink points by simply excluding the bunching window from the regression as described above. Intuitively, our approach provides an upper bound estimate for excess bunching at kink points.

more disposable income by earning (marginally) less. We define the share of individuals in the dominated region as $a = c^D/\hat{c}^D$, where c^D is the observed number of individuals in the dominated region $(k, k + D)$, and $\hat{c}^D$ is the counterfactual estimate of the individuals within the same region. D denotes the upper limit of the dominated region.

Similarly as in Chetty et al. (2011) and Kleven and Waseem (2013), the standard errors for all the estimates are calculated using a bootstrap procedure. We generate a large number of earnings distributions by randomly resampling the residuals from equation (5.4.1), and generate a large number of new estimates of $\hat{c}_j$ based on the resampled distributions to evaluate variation in the estimates of interest. The standard errors for each estimate ($\hat{b}(k)$ and $\hat{e}(k)$ for kinks, and $\hat{b}_E(k)$, $\hat{\delta}_H$, $\hat{a}$ and $\hat{e}(k)$ for notches) are defined as the standard deviation in the distribution of the estimate.

5.4.2. Data. We use panel data on all working-aged individuals (15-70 years) living in Finland in 1999-2011. The data set is based on the Finnish Longitudinal Employer-Employee Data (FLEED). To this data we have linked a variety of essential register-based variables, such as detailed tax register data for 1999-2011, and information on students and the study subsidy program for 1999-2010. With this data we can reliably and accurately analyze local changes in incentives among various subgroups of taxpayers. To analyze self-employed individuals, we use panel data on all main owners of Finnish businesses from 1999-2010, provided by the Finnish Tax Administration.

Table A3 in the Appendix presents the key summary statistics for all taxpayers. Table A4 shows the summary statistics for students. The average gross income of students excluding the study subsidy is 7,600€ per year. This implies that many students have part-time or full-time jobs during their studies and breaks between semesters, which is very typical among Finnish university students. Finally, Table A5 presents the summary statistics for self-employed individuals, including key firm-level characteristics.

5.5. Results

5.5.1. Baseline results. This section presents the overall results on bunching at MTR kink points and the study subsidy notch. We characterize the role and significance of frictions in the following sections.

Marginal tax rate kink points. First, we present taxable income distributions around different MTR kink points for all taxpayers. The figures plot the observed income distributions and counterfactual distributions relative to each MTR kink point in bins of 100€ in the range of +/- 5000€ from the kink. The figures denote the excess mass estimates (with standard errors), and the implied elasticity estimates based on observed excess bunching.

In each graph, the kink point is marked with a dashed vertical line. The excluded counterfactual region (the bunching window) is marked with solid vertical lines. In each graph, the bunching window is +/- 7 bins from the kink. The counterfactual density is estimated using a 7th-order polynomial function. Our results are not sensitive to the choice of the bunching window or the order of the polynomial.

Figure 5 presents the income distributions around different kink points of the central government income tax rate schedule for all taxpayers. The Figure illustrates bunching at the first, second, third and last kink point using pooled data for the years 1999-2011. As shown in Table A1 in the Appendix, the number of kink points has decreased from 6 to 4 in the period we study. Throughout the study, the first MTR kink point always includes the threshold where the central government income tax rate first applies. The other kink points in Figure 5 correspond to the kink points still existing after 2007.

The Figure shows that there is no bunching at the marginal tax rate kink points in Finland. The only conceivable exception might be the second kink. However, the second kink is likely to produce upward-biased excess bunching because of the locally hollow shape of the income distribution around the kink. Consequently, the elasticity estimates are zero or very close to zero at all MTR kink points.

In terms of the size of the tax rate change and the characteristics of the Finnish tax system, we should in particular find excess bunching at the first and last kink point of the tax schedule. However, there seem to be no significant responses at these kink points. The income distribution around the first kink point does not include individuals with means-tested social benefits, such as unemployment insurance payments and housing benefits. These benefits are regarded as taxable income, and they tend to cluster at certain income levels, causing a lot of noise in the low end of the income distribution. Importantly, there is no significant bunching at the first MTR kink even if we include individuals with these taxable benefits.

The result of no bunching also holds for all other central government kink points that are not shown in Figure 5. Also, there is no significant bunching at any kink point in any separate year. Thus there is no increase in excess bunching over time, and no differences in responses to kinks of different sizes. In addition to central government taxation, we find no bunching at the MTR kink points associated with graduated tax credits or allowances, including the municipal earned income tax allowance.

The result of no bunching at MTR kink points in Figure 5 indicates that marginal tax rates do not induce local behavioral responses. This could be explained by both the low underlying (local) tax elasticity and various behavioral frictions. It might be that the relatively small changes in incentives induce no behavioral responses, even in the absence of frictions. However, it might be that taxpayers cannot adjust their reported income or working hours with reasonable costs. Finally, it might be that taxpayers do not know or understand the details of the MTR schedule. We study these hypotheses by utilizing students and self-employed individuals as example groups in the next section.

Study subsidy notch. Next, we study behavioral responses around the notch points of the study subsidy system among Finnish university students. Figure 6 shows the

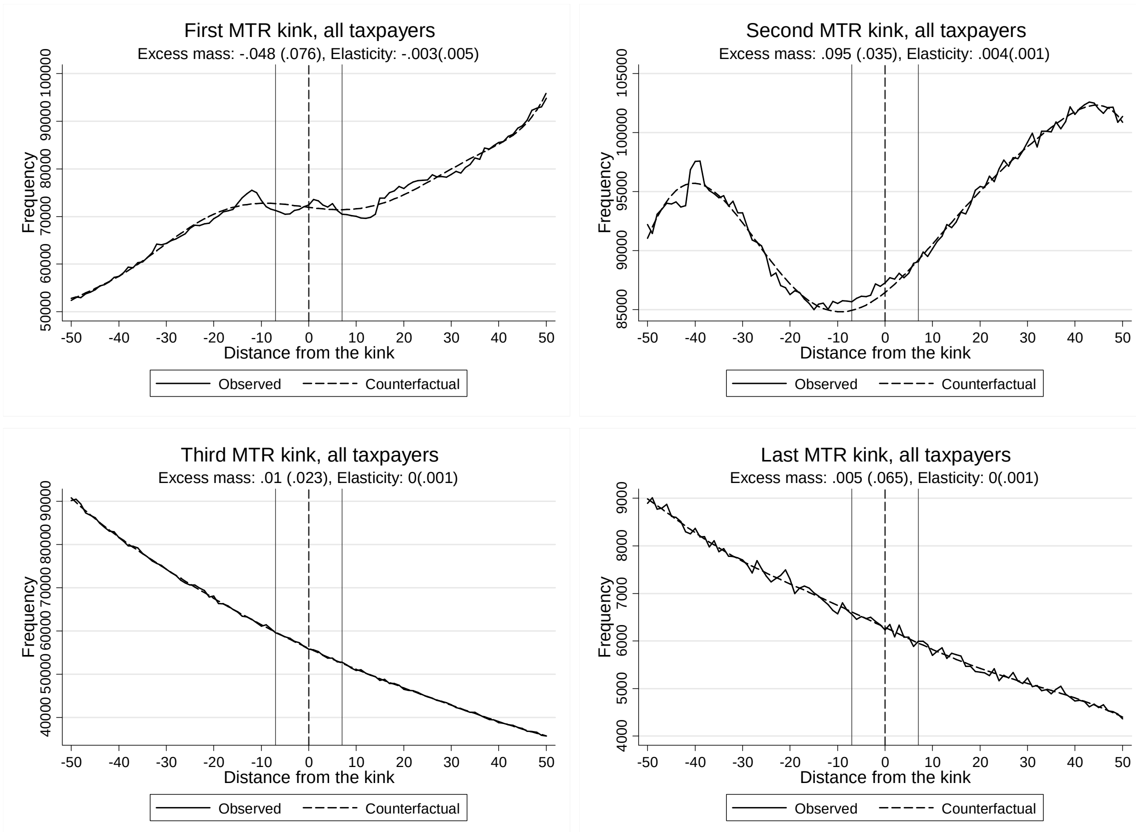


FIGURE 5. Income distributions around MTR kink points, 1999-2011

gross income distribution around the notch point (relative to the notch in bins of 100€ in the range of +/- 5000€ from the notch). The Figure presents the distribution of all students (left-hand side) and students with the default number of 9 study subsidy months (right-hand side) in 1999-2010. In the Figure, the dashed vertical line denotes the notch point above which a student loses one month of the subsidy. The solid vertical lines denote the excluded range (see Section 5.4 for details on defining the upper limit of the excluded range). The dash-point vertical line above the notch shows the upper limit of the dominated region.

The figure also includes the estimates and standard errors for the excess mass at the notch, the share of individuals in the dominated region, and the upper limit of the counterfactual and Δz. In each figure the counterfactual density is estimated using a 7th-order polynomial function. Our main conclusions are not very sensitive to this choice, although the point estimates vary somewhat with different choices of the degree of the polynomial.

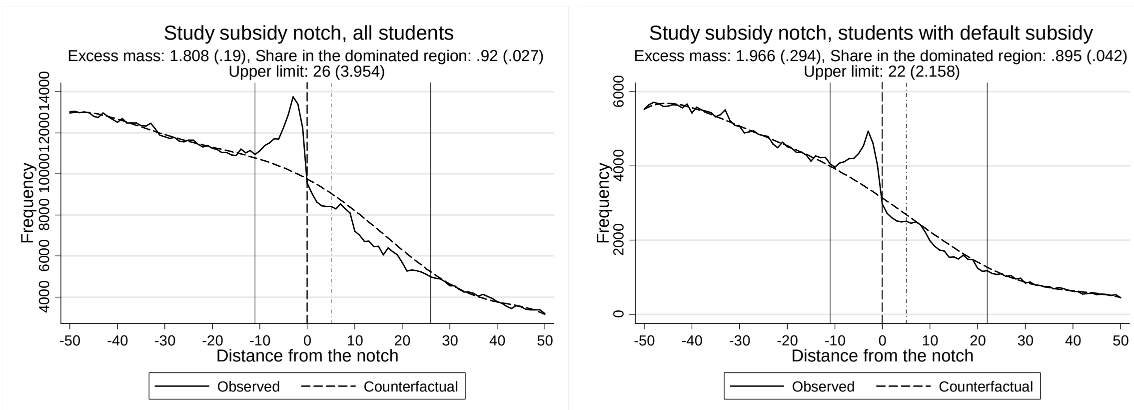


FIGURE 6. Bunching at the study subsidy notch, 1999-2010

Figure 6 indicates a clear and statistically significant excess mass on the left of the notch for both all students (1.8) and students with the default subsidy (2.0). This indicates that students are both aware of the notch and respond to the strong incentives created by it. However, the implied earnings elasticities are rather low, 0.083 (0.019) and 0.065 (0.007) for all students and students with 9 subsidy months, respectively (standard errors in parenthesis).[11] Thus even though excess bunching is evident and notable earnings responses occur ($\triangle z$ is around 15% of disposable income at the notch), the observed elasticities are still small. This stems from the fact that the changes in

[11]Earnings elasticity for all students is calculated using the average number of study subsidy months (7). All elasticities at study subsidy notches are calculated using the SISU microsimulation model and the average number of subsidy months. We thank Markus Paasiniemi for research assistance on calculating the elasticities.

incentives are also very distinctive, as notches induce very high implicit marginal tax rates above the income limit.[12]

5.5.2. Inability to respond.

Students. Figure 6 implies that students are aware of the incentives and respond to the notch created by the income limit of the study subsidy program. However, the Figure also suggests that students cannot affect their working hours or reported income very precisely. First, the excess mass below the notch is rather diffuse. This indicates that it is difficult for students to control or predict annual income very precisely. Second, there is an economically and statistically significant mass of students at the strictly dominated region above the notch where students can increase their net income by lowering their gross income. This indicates that in addition to the inability to respond, some students might not be aware of the notch, at least when exceeding the income limit for the first time.

To study the inability to respond, we first divide students into two groups based on the number of years they have studied: students with under three study years and students with more than (or equal to) three study years.[13] It is likely that the ability to adjust and predict annual income improves over time, therefore inducing the inability friction to decrease in the course of the study years. If the inability to respond decreases over time, we would expect that students with more study years bunch more actively, and that less students are located in the dominated region. In contrast, we have no

[12] In addition, implicit marginal tax rates remain relatively high (>50%) even further away above the notch, as an extra month of the subsidy is reclaimed after additional 1,010€ above the income limit (1,310€ after 2008). Thus, the effective tax schedule for students inherently includes multiple notches. However, we only observe significant bunching at the first notch, which justifies the analysis of the first notch only. The analysis of the first notch is also rationalized by the fact that students can alter the number of study subsidy months during the year.

[13] In order to eliminate the effect of dropouts and graduates, we include only students who also study in the next year.

clear reason to assume that willingness to respond would be different *locally* around the notch point for students with more or less study years.

Figure 7 weakly supports the above hypothesis. There is more excess bunching for more senior students, but the difference is not statistically significant.[14] Also, the share of individuals in the dominated region is practically unaffected. This suggests that a notable fraction of students are (still) not able to respond by adjusting their working hours. In addition, it might be that students with more experience of the study subsidy program are more aware of its details, and thus would respond more prominently. We study awareness of the study subsidy rules in more detail in the next subsection.

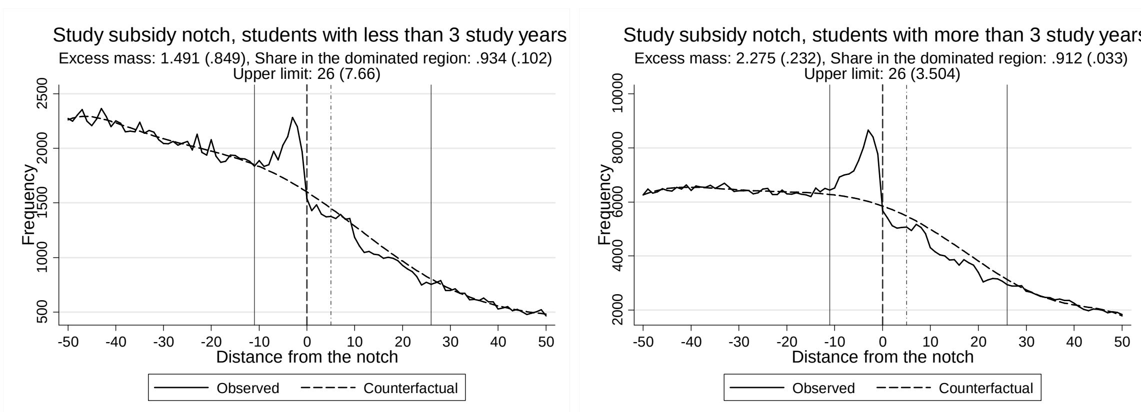


FIGURE 7. Bunching at the study subsidy notch: Students with more or less than 3 study years, 1999-2010

Next, we compare the responses of students around the study subsidy notch and the MTR kink points. There is a striking difference between bunching at notches and bunching at MTR kinks. Figure 8 shows the income distributions around MTR kink points for current students (first kink), university graduates (last kink) and students

[14]The earnings elasticities are 0.083 (0.035) and 0.083 (0.018) for students with less or more than three study years, respectively. Both elasticities are calculated using the average number of study subsidy months.

who previously bunched at the study subsidy notch (first kink). For all of these groups we find no significant bunching at any MTR kink point in any year.

Even though students are clearly responding to large incentives induced by the notch, they do not respond to smaller incentives created by MTR kinks. For current students this cannot be explained by an inability to respond to *any* local incentives, as we observe similar or even the same individuals bunching at income notches. In other words, there is no fundamental reason to assume that students are less able to affect their labor supply around the MTR kink compared to the study subsidy notch. Nevertheless, this result does not indicate that students would not respond to MTR kinks of *any* size. Larger changes in the MTR might induce larger observed behavioral responses, as with larger kinks it becomes more profitable to adjust labor supply (see Chetty et al. 2011, and Chetty 2012). However, in addition to the size of the incentive, the underlying elasticity and the inability frictions, it might be that the MTR schedule is too obscure for many students.

To further characterize the effect of the size of the incentive and the inability friction, we divide students into two groups by the number of study subsidy months they apply for. The default number of study subsidy months is 9, which students receive every year while studying if they do not wish to apply for additional subsidy months (up to 12), or decrease the number of subsidy months. There is a different income limit associated with each number of subsidy months which is inversely related to the number of months. In other words, less annual study subsidy gives a higher income limit. Table A2 in the Appendix shows the income limits for different numbers of study subsidy months and the relative loss incurred when the income limit is exceeded (both before and after the reform of 2008).

We analyze students with 2-5 or 6-8 study subsidy months. Both of these groups have made an active decision to opt out from the default option. Therefore, the difference between these groups is that the size of the tax incentive at the notch is weaker for

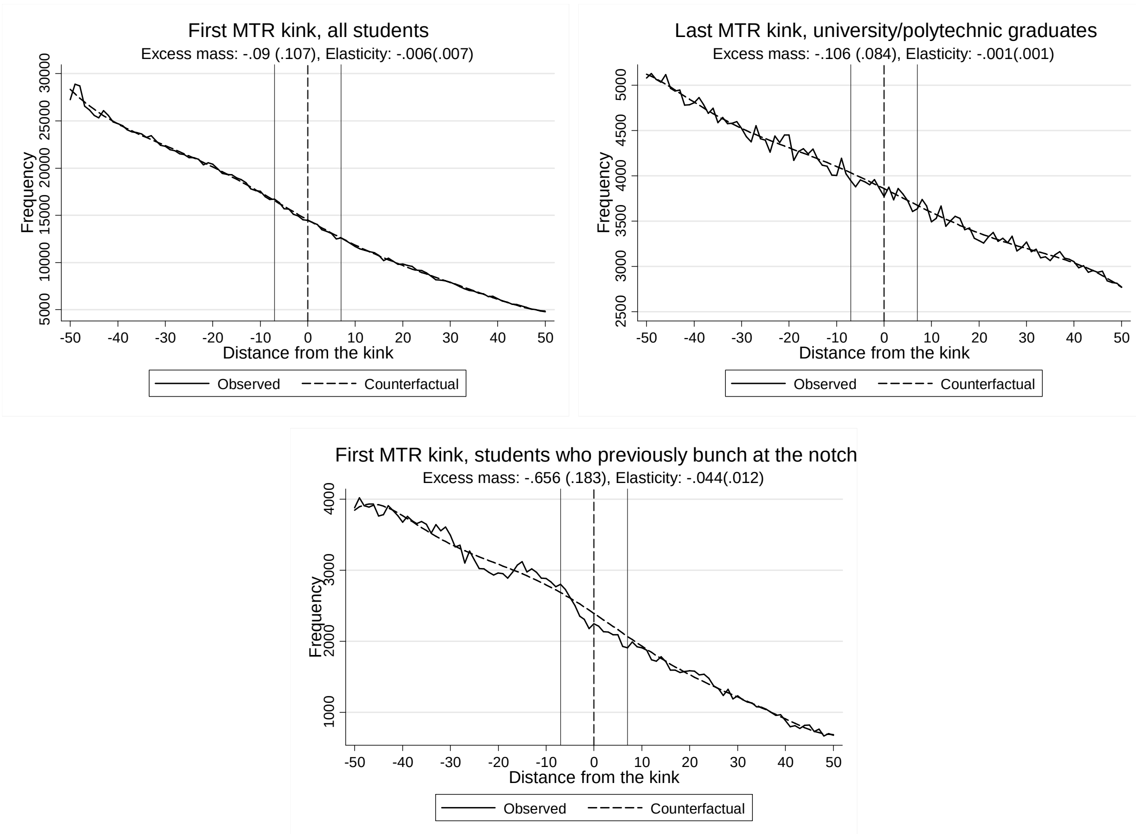


FIGURE 8. Bunching at MTR kink points: Current students, graduates and students who bunched at the study subsidy notch, 1999-2010

the group that had only a few months of the study subsidy, compared to the group that applied for more subsidy months. Importantly, both groups still lose disposable income if they earn marginally over the limit. In order to eliminate the possible peculiar effects of dropouts and graduates, we only include students who also study in the next year.

Figure 9 presents the bunching evidence at the study subsidy notch for students with different numbers of study subsidy months. Students with 2-5 subsidy months respond moderately, whereas for students with 6-8 months there seems to be a clear response at the notch. The observed elasticities are 0.07 (0.034) and 0.125 (0.042) for students with 2-5 and 6-8 subsidy months, respectively. However, it should be noted that the non-monotonic shape of the income distribution around the notch for students with 6-8 study subsidy months is not ideal when applying the bunching method, and there is large variation in the estimate of the counterfactual density when the degree of the polynomial is varied. Therefore, the estimates need to be interpreted with caution. However, the Figure clearly illustrates the notable behavioral response to the notch.

These findings support the notion that the relative strength of tax incentives matters, which corresponds to larger responses for larger tax incentives. This result also suggests that inability to respond matters. When inability to respond occurs, we should find more observed bunching at larger local changes in incentives. Furthermore, we have no explicit reason to assume that the two groups differ in awareness, especially as both groups have actively deviated from the default option. However, it is a possible that the inability to respond increases along with income for some students (i.e. along with more permanent or more highly paid jobs and more regular working hours). This would also indicate lower bunching when the income limit is higher.

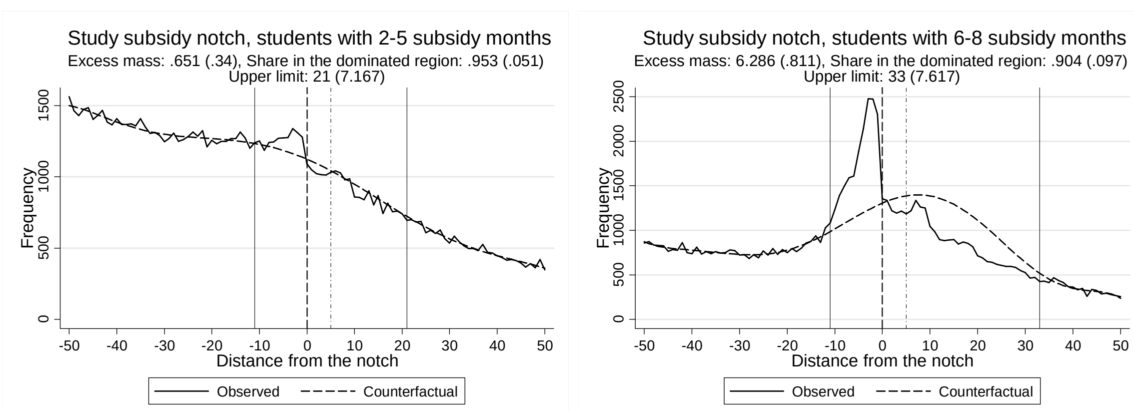


FIGURE 9. Bunching at the study subsidy notch: Students with different number of study subsidy months, 1999-2010

Self-employed individuals. Next we study the behavior of self-employed individuals. In this study, the self-employed include sole proprietors and partners of partnership firms (all non-corporate entrepreneurs in Finland). In many previous studies, the self-employed bunch much more actively at MTR kink points than regular wage earners. Saez (2010) finds clear and significant excess bunching for the self-employed in the US, Chetty et al. (2011) in Denmark, and Bastani and Selin (2014) in Sweden. One explanation for this finding is that entrepreneurs can more easily affect their labor supply and effort. Also, self-employed individuals have more opportunities to adjust their reported incomes. Combining the ability to respond and the large observed changes in behavior, earlier results indicate that entrepreneurs are aware of the incentives around kink points and respond to them actively.

In contrast to earlier findings form other Nordic countries (Chetty et al. 2011, and Bastani and Selin 2014), we do not find sharp and distinctive excess bunching for the self-employed at the kink points of the earned income tax rate schedule. This result holds for all kink points and, for example, for self-employed individuals in different industries and counties. As an example, the two graphs in Figure 10 show the income distributions around the first and last kink point for all Finnish self-employed individuals in 1999-2011. For the first kink point, the excess bunch estimate is statistically significant. However, the relatively large overall variation in the income distribution makes the evaluation sensitive to choices regarding the bunching window and the degree of the polynomial. For example, when decreasing the order of polynomial of the counterfactual density from 7 to 5, the excess mass at the first kink decreases to 0.083 (0.08). Thus we can conclude that MTR kink points do not induce robust and significant local responses among the self-employed individuals in Finland.

In order to assess the ability to affect reported income among self-employed individuals, we study bunching at round numbers in terms of personal *gross* earned income.[15]

[15]The connection between gross earned income and taxable earned income is that taxable earned income is calculated by deducting individual tax allowances and tax deductions from gross earned income.

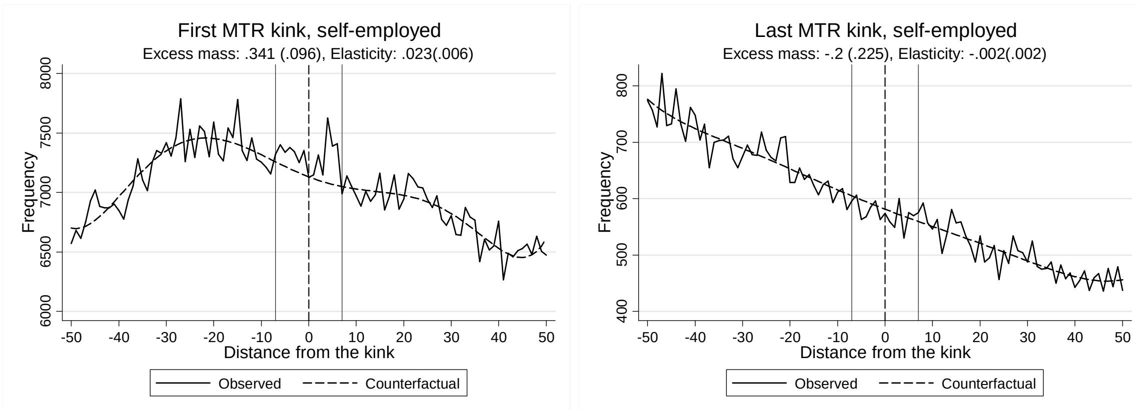


FIGURE 10. Bunching at MTR kink points: Self-employed individuals, 1999-2011

Significant round-number bunching would indicate that at least some self-employed individuals can affect the gross income they report to the tax authorities. Thus if we find evidence of self-employed bunching at convenient round numbers, it suggests that entrepreneurs have some ability to affect their location in the income distribution.

Figure 11 shows that the self-employed bunch sharply at gross earned income of 10,000€ and 20,000€.[16] This behavior also occurs at larger round numbers (30k, 40k, 50k etc.) as well as at round numbers of Finnish marks (100k, 200k, 300k) before the implementation of the euro in 2002. Furthermore, we find that significant round-number bunching never occurs among regular wage earners.

Combining the evidence from Figures 10 and 11 suggests that inability to respond might not explain the no-bunching result at kink points for self-employed individuals. In contrast to different responses among students at notches and MTR kink points which both induce changes in tax incentives, there is no "payoff" from bunching at round numbers. Based on standard economic theory, locating just below the kink point instead of a round number would induce some (small) utility gain. Overall, it seems

[16]As round-number bunching is very sharp by nature, we use a bunching window of +/- 100€ on both sides when defining excess bunching at round numbers.

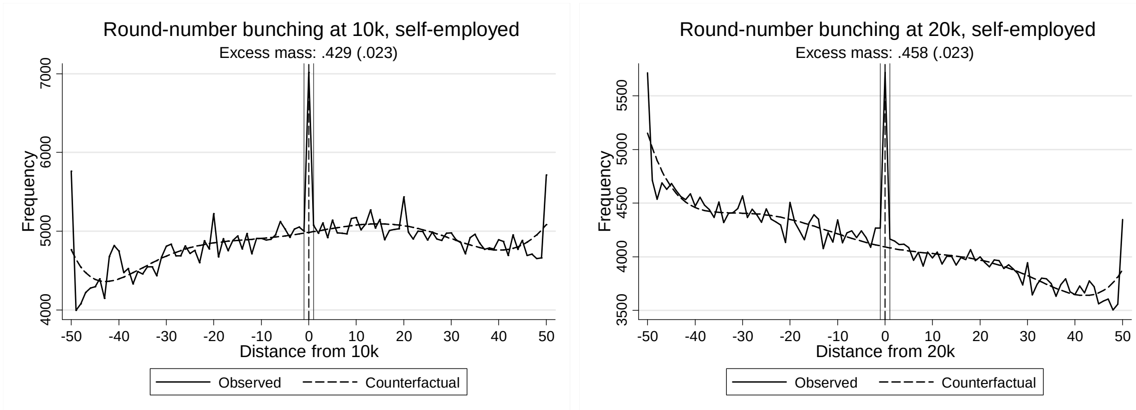


FIGURE 11. Round-number bunching, gross earned income: Self-employed individuals, 2002-2011

that either unawareness of the details of the MTR schedule or low underlying elasticity at smaller changes in incentives might be driving the results, but we cannot precisely separate between the two mechanisms. However, we discuss the potential issues related to awareness of the tax rules in the next subsection.

5.5.3. Unawareness.

Students. To study the unawareness friction, we first analyze students who previously located themselves in the dominated region just above the notch. In this region, students could earn more disposable income by earning less gross income. In addition, students who exceed the income limit receive a letter from the Social Security Institution which states that (at least) one month of the subsidy needs to be paid back (with 15% interest). Thus for the students who are just over the income limit, there are both large incentives to adjust behavior in the future as well as increased awareness of the incentives and the existence of the income limit due to the letter they receive.

Figure 12 shows the income distribution around the notch for those students who were located in the dominated region in any of the three previous years. The Figure shows that students bunch actively at the notch after locating in the dominated region before. The elasticity estimate at the notch for this group is 0.112 (0.051). However, a notable share of individuals still fail to optimize and are located in the dominated region also in future years. This suggests that inability to respond is a prominent feature even with large incentives and when awareness is generally increased. However, the non-monotonic shape of the income distribution around the notch for this particular group induces notable variation in the estimates, which thus need to be interpreted with caution.

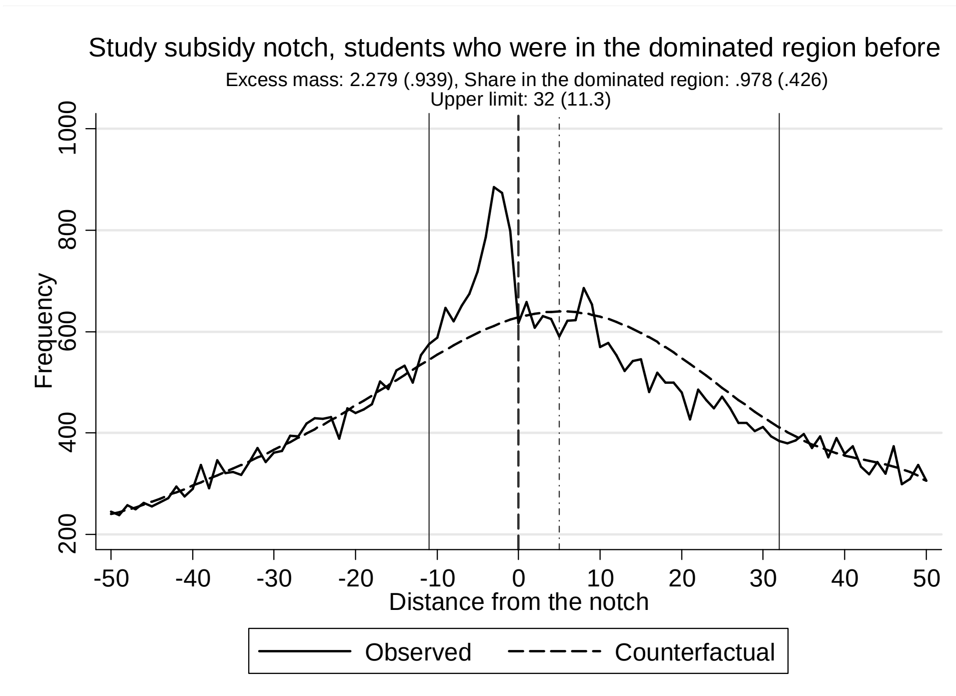


FIGURE 12. Bunching at the study subsidy notch: Students who were in previous years in the dominated region ($t-1$, $t-2$ or $t-3$,), 1999-2010

Second, we characterize the effects of the study subsidy reform of 2008. In the reform, the income limits increased by approximately 30%. For the default 9 months of the subsidy, the reform increased the income limit from 9,260€ to 12,070€. At the same time, the overall study subsidy was increased from 461€ to 500€ per month.

First, Figure 13 shows the income distribution around the study subsidy notch point for all students before (1999-2007) and after (2008-2010) the reform. From the Figure we can see that there is excess mass at the notch point both after and before the reform. This indicates that students are aware of the income limit and its changes. Furthermore, the excess mass at the old income limit disappears immediately after the reform. Figure A2 in the Appendix shows the income distribution around the old income limit for all students in 2008-2010.

However, compared to 1999-2007, the excess mass is smaller after the reform. This tentatively suggests that either students are not as aware of the new notch point as they are of the old one, or it takes time to learn the new rules. Supporting this, the elasticity estimate from before the reform (0.092(0.011)) is somewhat larger than the observed elasticity at the notch after 2008 (0.054 (0.013)).

Gradual learning and a gradual increase in awareness over time are also tentatively supported when we study how students respond to the notch in different years before 2008. The study subsidy system with the current income limits was introduced in 1998. Thus the system was still relatively new in the late 1990s and the early 2000s. Indeed, Figure A3 in the Appendix shows that the excess mass has increased over time when we compare the years 1999-2001, 2002-2004 and 2005-2007, but the differences are in general not statistically significant. Importantly, we have no general reason to assume that the ability to respond to the notch would be different between different years. Furthermore, the income limits have remained unchanged over this time period, with the exception of minor adjustments after the implementation of the euro in 2002.

Nevertheless, when comparing excess bunching at the notch before and after the reform of 2008, it might be that the observed behavioral responses are different at higher income levels due to different inability frictions. We discussed this issue in the previous section when characterizing the effect of different study subsidy months (Figure 9).

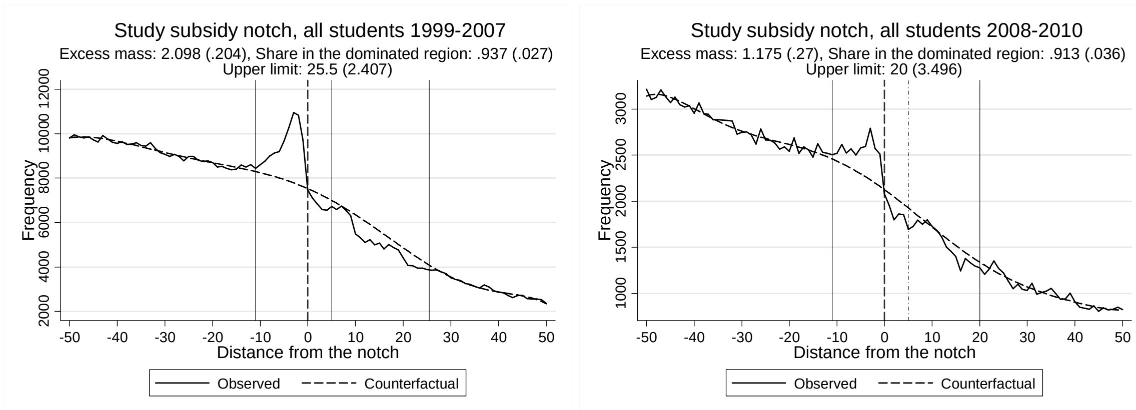


FIGURE 13. Bunching at the study subsidy notch: Before the reform 1999-2007 (left-hand side) and after the reform 2008-2010 (right-hand side)

To characterize the effect of the change in the income limit more closely, we analyze the income distribution around the notch point after the reform of 2008 among those students who bunched at the income notch in any of the years 2005-2007 before the reform. Figure 14 shows that students who bunched at the lower income limit before also bunch actively when the income limit is increased. The observed earnings elasticity at the notch for this group is 0.067 (0.027). One implication of this result is that students are in general aware of the income limit, and actively adjust their behavior to changes in study subsidy regulations. However, from the Figure we can again see that a large number of students are located in the dominated region above the notch. Thus this behavior occurs even among the same individuals who were before located just below the notch. This lends more support to the earlier conclusion that inability frictions related to real labor supply decisions play an important role even with large behavioral incentives which taxpayers are in general aware of.

Self-employed individuals. Our finding of no distinctive bunching behavior among self-employed individuals differs from recent studies in other Nordic countries. In contrast to Sweden (Bastani and Selin 2014) and Denmark (Chetty et al. 2011), we do not

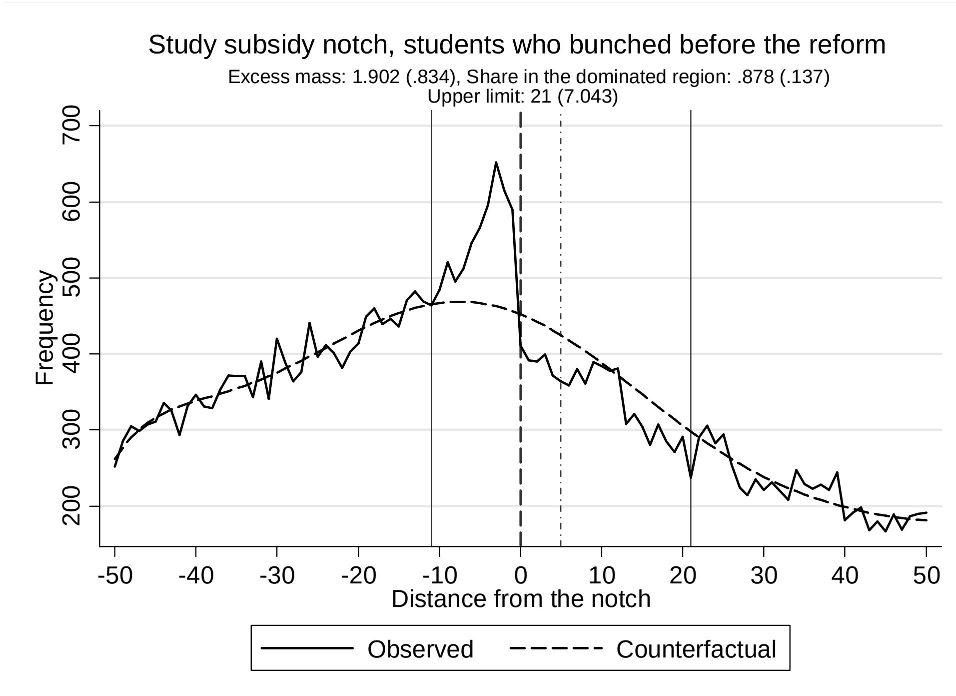


FIGURE 14. Bunching at the study subsidy notch: Students who bunched before the reform (2005-2007), 2008-2010

find sharp bunching at MTR kink points for the self-employed. This is especially intriguing since, for example, Finland and Sweden have rather similar overall tax systems and other institutional features.

One issue explaining the different responses could be the size of the local change in incentives. However, the relative size of the kink points is rather similar in Denmark and Sweden compared to the largest kink points in Finland. For Denmark, the first jump in the MTR causes (approximately) a 11% fall in the net-of-tax rate, and the second kink 30% (Chetty et al. 2011). For Sweden, the first kink point decreases net-of-tax rates by around 34%, and the second kink by 10% (Bastani and Selin 2014). In Finland, the first kink point causes a decrease of 22-53% in net-of-tax rates, and the last kink decreases net-of-tax rates by 9-16% (depending on the year in question). Thus there is no dramatic difference between the size of the change in incentives.

However, one important difference could be the number of kink points and the salience of the tax schedule. Compared to Sweden and Denmark, the Finnish income tax schedule includes more kink points. In Sweden and Denmark there are only two distinctive jumps in marginal tax rates (Chetty et al. 2011 and Bastani and Selin 2014). In Finland, there are 4-6 kink points in the central government tax schedule in the period we study. At least to some extent, the small number of kinks might make the tax schedule more transparent in Denmark and Sweden compared to the Finnish tax schedule. Thus assuming that self-employed individuals have similar underlying preferences in Finland and Sweden, the difference of the response might be due to lower awareness and transparency of the tax system in Finland. However, as discussed in Bastani and Selin (2014), there might be relevant differences between the tax systems with respect to the ability to affect reported income through, for example, tax avoidance or tax evasion. This might also have a notable effect on observed bunching.

Overall, as noted before, the negligible bunching at MTR kink points among the self-employed might be due to different frictions, and it is challenging to distinguish between different mechanisms. In future work we aim at a more careful distinction between different frictions by analyzing different subgroups of self-employed individuals, including the owners of privately held corporations.

5.6. Implications and conclusions

In this paper we study the role of individual preferences and different optimization frictions in explaining taxpayer responses to tax incentives. We compare the behavioral effects induced by various tax incentives within the same institutions and subpopulations. We apply the bunching method to kinks and notches of different strength in order to produce clear, comparable and visually convincing evidence.

We find no bunching at kink points of the marginal income tax rate schedule. This result holds for any subgroup we study, and for kink points of any size in any year. Intriguingly, the result of negligible bunching holds for the self-employed, although earlier studies find notable and sharp bunching behavior among this group (see e.g. Chetty et al. 2011 for Denmark, and Bastani and Selin 2014 for Sweden). We present clear evidence that the self-employed in Finland are able to control their reported income, as they actively bunch at round numbers of reported gross income. This evidence suggests that the inability to respond does not fully explain the lack of bunching at kink points.

We find that students bunch actively at the income notch induced by the study subsidy system. Since even the same students who bunch at the notch do not bunch at marginal tax rate kink points, we focus on explaining why the behavioral responses differ. First, notches induce considerably larger incentives, and should in theory create more significant bunching behavior. However, even with a relatively small elasticity we should observe at least some bunching at kink points, which we can, however, rule out with our large data set. Therefore, different frictions related to these different institutions might explain the differences.

One important factor is the visibility of tax incentives and awareness of tax rules. The study subsidy, like many other means-tested benefits, needs to be applied for. Moreover, the income limit is relatively salient, and we find that students clearly noticed recent changes in the income limits. However, the marginal tax rate schedule is presumably less salient, and unawareness might partly explain the non-response at kink points. Importantly, for students we can hold the inability to respond constant, as they are in principle as able to bunch at the notch as they are at the kink point.

We observe that many students are located in the dominated region above the notch. This holds even for the same students who optimized and located at the notch in previous years. This evidence suggests that the inability to respond related to labor market rigidities plays an important role. The majority of students who work are wage earners, and altering reported income is not easy for them in the Finnish context. If students want to affect their annual income, they can, for example, either find another job or stop working abruptly at a certain point of time during the year when the income limit is reached. These types of decisions might have relatively large costs, which translates into large frictions.

Our results so far point to the following implications: First, behavioral responses to kink points might be at least partly attenuated because of frictions related to unawareness. Unaware individuals do not respond to tax incentives, and it is possible that they will not respond in the future either, at least without a considerable change in awareness of the regulations. Then the long-run elasticity would be close to the observed behavioral responses locally at kink points. This would indicate near zero local elasticity of taxable income, at least for small changes in incentives.

Second, when individuals are aware of incentives that are strong enough, we observe clear behavioral responses. These stronger responses are associated with changes in average tax rates. However, our empirical evidence suggests that these responses are attenuated by the inability to respond, and that the underlying long-run structural elasticity might be somewhat larger than the observed one if taxpayers can adjust their behavior over time.

In summary, different reasons for responding and not responding to tax incentives can entail different welfare and policy implications. Therefore, it is important to distinguish what types of taxes create behavioral responses, and what types of frictions affect individual behavior in different institutional contexts. In future work we aim at a more careful distinction between different frictions by using changes in policy rules and a more careful divided sample analysis.

Bibliography

[1] Bastani, S. and H. Selin. (2014). Bunching and non-bunching at kink points of the Swedish tax schedule. Journal of Public Economics, 109: 36–49.

[2] Blundell, R. and A. Shepard. (2012). Employment, hours of work and the optimal taxation of low income families. Review of Economic Studies, 79: 481–510.

[3] Blundell, R. and H. Hoynes. (2004). Has in-work benefit program helped the labor market?, in Blundell, R. D. Card and R. Freeman (eds.)., Seeking a Premier Economy: The Economic Effects of British Economic Reforms, 1980-2000. Chicago: Chicago University Press, 2004.

[4] Chetty, R., J. Friedman and E. Saez. (2013). Using differences in knowledge across neighborhoods to uncover the impacts of the EITC on earnings. American Economic Review, 103(7): 2683–2721.

[5] Chetty, R. and E. Saez. (2013). Teaching the tax code: Earnings responses to an experiment with EITC recipients. American Economic Journal: Applied Economics, 5(1): 1–31.

[6] Chetty, R. (2012). Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply. Econometrica, 80(3): 969–1018.

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

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

[9] Chetty R., A. Looney and K. Kroft. (2007). Salience and taxation: Theory and evidence. NBER Working Paper 13330.

[10] Feldstein, M. (1999). Tax avoidance and the deadweight loss of the income tax. Review of Economics and Statistics, 81: 674–680.

[11] Gruber, J. and D. Wise. (1998). Social security and retirement: An international comparison. The American Economic Review, 88(2) Papers and Proceedings of the Hundred and Tenth Annual Meeting of the American Economic Association: 158–163.

[12] Kleven, H. and M. Waseem. (2013). Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan. Quarterly Journal of Economics, 128: 699–723.

[13] Liebman, J. and R. Zeckhauser. (2004). Schmeduling. Manuscript.

[14] Manoli, D. and A. Weber. (2011). Nonparametric evidence on the effects of financial incentives on retirement decisions. NBER Working Paper 17320.

[15] Meghir, C. and D. Phillips. (2010). Labour supply and taxes. in S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gummie, P. Johnson, G. Myles and J. Poterba (eds.). Dimensions of tax design: The Mirrlees review. Oxford, Oxford University Press, 2010, 202–274.

[16] Reck, D. (2014). Taxes and mistakes: what is in a sufficient statistic? Manuscript.

[17] Saez, E. (2010). Do taxpayers bunch at kink points? American Economic Journal: Economic Policy, 2(3): 180–212.

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

[19] Sallee, M. and J. Slemrod. (2012). Car notches: Strategic automaker responses to fuel economy policy. Journal of Public Economics, 96: 981–999.

[20] Slemrod, J. (2010). Buenas notches: Lines and notches in tax system design. Working Paper.

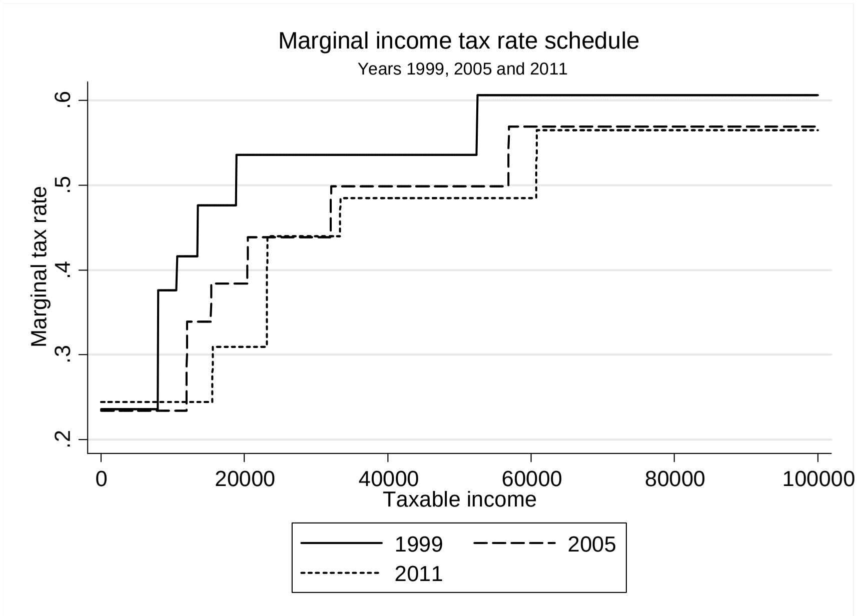
[21] Yelowitz, A. (1995). The Medicaid notch, labor supply and welfare participation: Evidence from eligibility expansions. Quarterly Journal of Economics, 110: 909–939.

Appendix

Year	Taxable income (in euros)	Marginal tax rate
1999	7,905-10,596	5,5
	10,596-13,455	15,5
	13,455-18,837	19,5
	18,837-29,601	25,5
	29,601-52,466	31,5
	52,461-	38
2000	8,006-10,697	5
	10,697-13,623	15
	13,623-19,005	19
	19,005-29,937	25
	29,937-52,979	31
	52,979-	37,5
2001	11,100-14,296	14
	14,296-19,678	18
	19,678-30,947	24
	30,947-54,661	30
	54,661-	37
2002	11,500-14,300	13
	14,300-19,700	17
	19,700-30,900	23
	30,900-54,700	29
	54,700-	36
2003	11,600-14,400	12
	14,400-20,000	16
	20,000-31,200	22
	31,200-55,200	28
	55,200-	35
2004	11,700-14,500	11
	14,500-20,200	15
	20,200-31,500	21
	31,500-55,800	27
	55,800-	34
2005	12,000-15,400	10,5
	15,400-20,500	15
	20,500-32,100	20,5
	32,100-56,900	26,5
	56,900-	33,5
2006	12,200-17,000	9
	17,000-20,000	14
	20,000-32,800	19,5
	32,800-58,200	25
	58,200-	32,5
2007	12,400-20,400	9
	20,400-33,400	19,5
	33,400-60,800	24
	60,800 -	32
2008	12,600-20,800	8,5
	20,800-34,000	19,0
	34,000-62,000	23,5
	62,000 -	31,5
2009	13,100-21,700	7
	21,700-35,300	18
	35,300-64,500	24
	64,500 -	30,5
2010	15,200-22,600	6,5
	22,600-36,800	17,5
	36,800-66,400	22,5
	66,400 -	30
2011	15,600-23,200	6,5
	23,200-37,800	17,5
	37,800-68,200	22,5
	68,200 -	30

Note: Finnish marks are converted to euros before 2002.

Table A1. Central government marginal income tax rates, 1999-2011



Note: Marginal tax rate includes central government income taxes, average municipal income taxes and average social security contributions.

FIGURE A1. Nominal marginal tax rates (MTR) on earned income, years 1999, 2005 and 2011

	Before 2008		*After 2008*	
Study subsidy months	Income limit	Relative income loss at the margin if income limit is exceeded	Income limit	Relative income loss at the margin if income limit is exceeded
1	17,340	3.1%	22,550	2.5%
2	16,330	3.2%	21,190	2.7%
3	15,320	3.5%	19,930	2.9%
4	14,310	3.7%	18,620	3.1%
5	13,300	4.0%	17,310	3.3%
6	12,290	4.3%	16,000	3.6%
7	11,280	4.7%	14,690	3.9%
8	10,270	5.2%	13,380	4.3%
9	9,260	5.7%	12,070	4.8%

Note: The relative loss from marginally exceeding the income limit is calculated using the full study subsidy (461 euros and 500 euros before and after 2008, respectively) plus 15% interest charged by the Social Insurance Institution.

TABLE A2. Income limits in the study subsidy system and the relative marginal loss if the income limit is exceeded (in proportion to gross income at the limit), before and after the reform of 2008 (academic years 2006/2007 and 2008/2009, respectively)

Variable	N	Mean	Std. Dev.
Taxable earned income	45,494,860	22,981	31,048.75
Gross earned income	45,494,615	25,520	31,986.62
Taxable capital income	45,494,860	1,803	50,947.13
Age	45,494,860	43.06	14.816
Female	45,494,860	0.498	.50
Size of the household	44,963,949	2.68	1.438

TABLE A3. Summary statistics, all taxpayers, 1999-2011

All students	N	Mean	Std. Dev.
Taxable income	3,970,775	6,115	5,072.32
Gross income (subject to income limit)	2,711,754	7,614	8,619.77
Age	3,980,502	23.30	5.093
Subsidy months	3,255,567	7.15	2.762
Income limit	3,249,902	11,730	3,206.64
Students with 9 months of study subsidy	N	Mean	SD
Taxable income	1,163,189	5,024	3,879.44
Gross income (subject to income limit)	708,525	5,587	6,492.77
Age	1,163,617	22.54	4.486
Subsidy months	1,163,617	9	0
Income limit	1,163,189	9,770	1,083.15

TABLE A4. Summary statistics, students, 1999-2010

Variable	N	Mean	Std. Dev.
Taxable earned income	3,351,466	25,601	26,860
Gross earned income	3,385,734	26,970	27,734.4
Capital income	2,236,182	7,096	561,620.2
Turnover (firm-level)	3445810	149,931	663,409.8
Net assets (firm-level)	2956521	15,211	217,001.8
No. of employees (firm-level)	2189383	.661	2.587

TABLE A5. Summary statistics, self-employed (sole proprietors and partners of partnership firms), 1999-2011

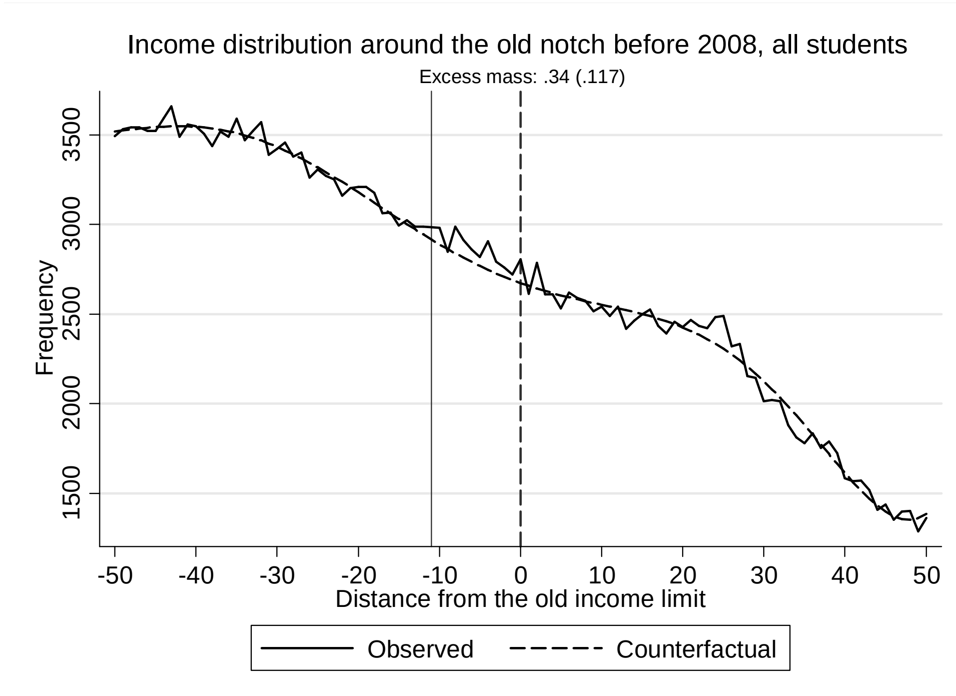


FIGURE A2. Income distribution around the old income limit before the reform of 2008, all students 2008-2010

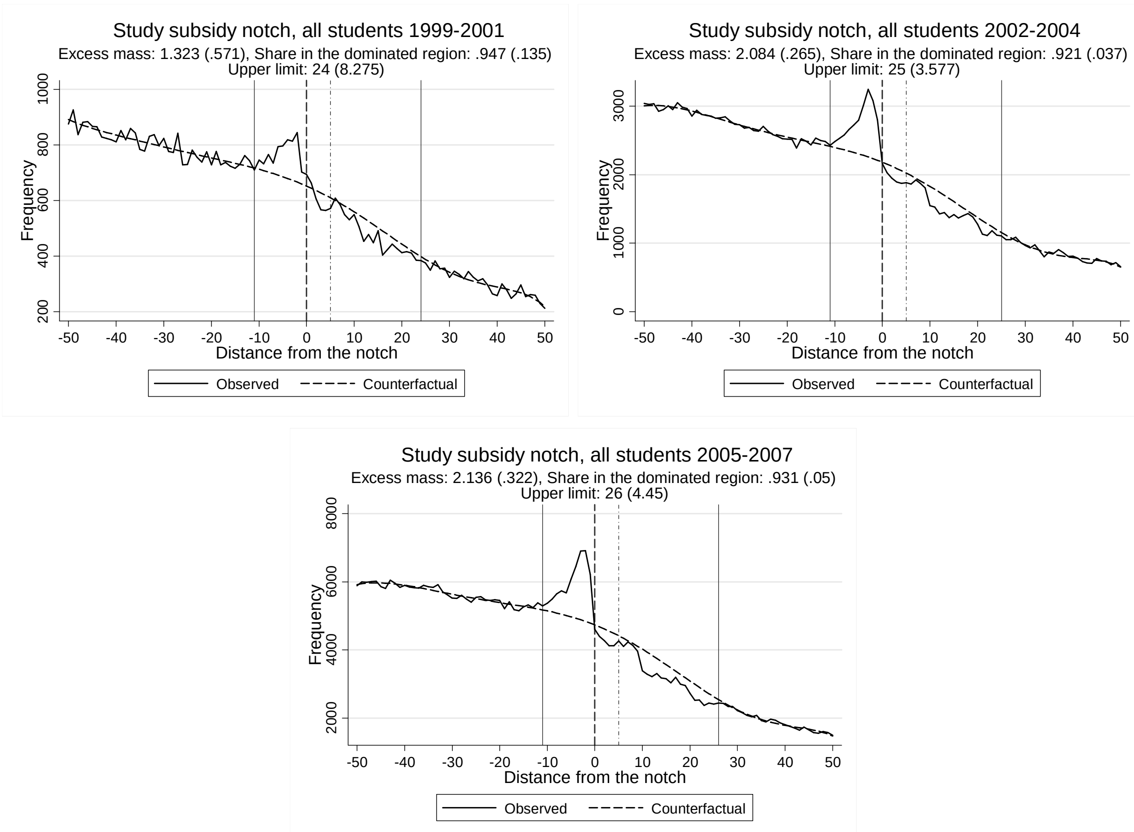


FIGURE A3. Bunching at the study subsidy notch in different years