The Dynamics of Labor Supply Responses to Tax Changes.

Interpretations and Comparisons of Results

from Microsimulation Models and Panel Data Approaches.

INAUGURAL-DISSERTATION

zur Erlangung des Grades
eines Doktors der Wirtschaftswissenschaft
(doctor rerum politicarum)
am Fachbereich Wirtschaftswissenschaft
der Freien Universität Berlin

vorgelegt von
Trine Engh Vattø, M.A.

Berlin, 2013

Erstgutachter:

Prof. Dr. Viktor Steiner, Freie Universität Berlin

Zweitgutachter:

Prof. Dr. Frank Fossen, Freie Universität Berlin

Datum der Disputation: 01.07.2013

Acknowledgments

I had the unique opportunity to write this doctoral thesis in Berlin with the support from Statistics Norway. I am grateful to have Viktor Steiner as my principle supervisor within the framework of the Berlin Doctoral Program of Economics and Management Science (BDPEMS). Since I enrolled in the PhD program in 2009, I have benefited immensely from a wide range of courses and seminars, and from my fellow doctoral students.

First of all, I am indebted to Viktor Steiner for supervising my thesis with valuable discussions and critical questions. I have learned a lot from his experience and insight in empirical research. I would also like to thank my second supervisor, Frank Fossen, for his valuable support and agreeing to evaluate my thesis. Further, I am much obliged to Peter Haan for providing me the opportunity of a guest position at the German Institute for Economic Research (DIW Berlin), and let me be a part of a motivating and enjoyable research environment.

This thesis would not have been possible without the support of Statistics Norway. I owe a special thanks to Thor Olav Thoresen who encouraged and guided me through all stages of my PhD-studies, and for the collaboration on Chapter 2. I would further like to thank Zhiyang Jia for his help and advice in tricky estimation issues, and for the collaboration on Chapter 4. Moreover, I am indebted to John Dagsvik for the time he spent teaching me and broadening my horizon. Financial support to work on this thesis from Statistics Norway and the Research Council of Norway was gratefully appreciated.

The weekly “Wirtschaftspolitisches Forschungsseminar” led by Viktor Steiner and Frank Fossen has been stimulating, and I am thankful for valuable comments from all participants while presenting my own research. I would in particular like to thank Clive Werdt for numerous discussions on common research topics and for commenting on earlier drafts of my papers and presentations. I am moreover thankful for a number of opportunities to present my research at international conferences and workshops.

Finally, I would like to thank friends and family for proofreading, help and encouragement. A special thanks goes to Michael Brandt for technical guidance and for being there for me while working on this thesis.

Contents

Chapter 1

Introduction

Labor supply responses to tax changes are a core issue in public economics, reflected by numerous estimates from different methodological approaches. Relationships between labor supply and taxes in a microeconomic and microeconometric perspective are often discussed based on two categories of research. The first line of research refers to estimation of structural labor supply models for simulation purposes. From observations of individuals' consumption and working hours choices, one can estimate parameters that reflect agents' preferences, which in turn can be used to simulate the effects of changes in the economic budget constraint such as changes in the tax schedule. A second main method is based on analyses of observations before and after a realized policy reform, comparing outcomes for groups affected or not affected by a particular change in the tax system. The elasticity of taxable income (ETI) is here a key concept which measures the response in taxable income to a change in the net-of-tax rate (defined as $1-$marginal tax rate). ETI represents a possible broader set of labor supply responses than the traditional focus on working hours.

Along both empirical methods (as well as in the theoretical literature) it is typically assumed that individuals react instantly to changes in the tax schedule, either at the time of implementation of a policy reform or already at the time when the policy reform is announced. However, there has been a growing awareness in the labor supply literature of explanations to why people tend to stick to their original choice of work. Recently, it has been suggested that adjustment or search costs and labor market frictions should not be neglected when analyzing microdata with relatively small changes in the tax schedule (Chetty, 2012; Chetty et al., 2011; Blundell et al., 2011). State dependence, defined as the causal effect of experiencing an event on preferences, prices or constraints relevant to future choices (Heckman, 1981), might represent another source of observed inertia in people's reactions to policy changes.

How optimization frictions and state dependence affect the dynamics of adjustment in labor supply to tax changes over time is a topic which has been generally overlooked in the literature (two exceptions are Holmlund and Söderström, 2011 and Haan, 2010). There are many possible reasons for slower responses in labor supply and consequently the need for a model framework which incorporates the time dimension. One argument is people's lack of information regarding changes in the tax schedule, which slowly improves after experiencing a change in disposable income. Other arguments are habit persistence and costs associated with changes in hours worked (Holmlund and Söderström, 2011). Moreover, when analyzing the broader set of labor supply responses through earnings, some responses such as increased effort, responsibility and human capital investment (or on the job training) might lead to a rise in earnings only after some time has passed. Following the argument in Slemrod (1992, 1995) one can think of a hierarchy of behavioral responses, in which timing of transactions, avoidance behavior and real decisions of individuals and firms can be ordered from the most to the least responsive. One reason for this conception might be that responses related to timing of transactions are immediate or apply already from the tax reform announcement, whereas the responses in labor supply might be considerably slower due to adjustment costs.

This dissertation contributes to the literature along two main avenues. First, I demonstrate how simulations from a static discrete choice model can be validated by panel data methods estimating the elasticity of taxable earnings (Chapter 2). Even though there are some examples of studies which discuss quasi-experimental evidence in relation to results from structural models, e.g. Blundell (2006), Eissa and Hoynes (2006) and Todd and Wolpin (2006), there has so far been less cross-checking of results from the ETI studies and structural labor supply model simulations. Second, I analyze the dynamics of labor supply responses to tax changes over time both with regards to panel data approaches (Chapter 3) and incorporated in a structural labor supply framework (Chapter 4).

I focus on responses in working hours and labor income for wage earners, and use Norwegian data from three main sources provided by Statistics Norway. First, for all analyses, I use Income Statistics for Persons and Families, which is detailed tax return data included with a wide range of individual characteristics such as education level and field of education for the complete Norwegian population. Second, to estimate the static discrete choice model I utilize cross-sectional data from the Norwegian Labor Force Survey, which gives detailed information about working hours for a representative sample covering about 1 percent of wage earners. And third, in order to estimate the intertemporal structural model, I use register data on working hours from the Norwegian Earnings Survey which covers all public sector employees and about 70 percent of private sector employees. A personal identification number makes it possible to track people over time and combine the different data sources.

In Chapter 2, I (joint work with Thor Olav Thoresen) show how the standard ETI methodology can be used to validate predictions from a structural labor supply model, by analyzing the Norwegian tax reform in 2006. The structural labor supply model discussed, following e.g. Dagsvik and Jia (2012), and Dagsvik et al. (2013), is related to the discrete model of Soest (1995). A similar model is available to Norwegian decision-makers through the model system LOTTE (Aasnes et al., 2007; Thoresen et al., 2010). Traditional methods in the ETI literature are used (following e.g. Gruber and Saez, 2002), utilizing the panel structure of data to obtain individual measures of income growth, and employing instrumental variable techniques to obtain measures of change in the net-of-tax rate. To facilitate comparison, I use the discrete choice model to simulate the effects of the 2006 tax reform on hours of work, and use predicted income levels to obtain a comparable estimate for income elasticity with respect to the net-of-tax rate. Both methods suggest small responses (elasticities below 0.1) in labor supply and earnings to the particular tax reform of consideration.

In Chapter 3, I analyze the dynamics of earnings responses to tax changes by exploiting substantial exogenous variation in the two-tier surtax schedule for labor income over a period of 14 years (1995-2008). I adopt the dynamic panel data framework by Holmlund and Söderström (2011), and compare with estimates from the conventional static panel approach by Gruber and Saez (2002). The estimated magnitude of the elasticities of earnings with respect to net-of-tax are modest, about 0.06 for the three-year conventional static panels, and about 0.06/0.12 in the short/long run for the dynamic specifications. I find that the long run responses to tax rate changes in labor supply are about twice as large as the short run responses due to a strong autoregressive effect in earnings. Supplementary conducted simulations suggest that up to 40 percent of the revenue loss due to tax rate cuts in the surtax schedule can be self-financed in the long run by inducing additional generated income.

In Chapter 4, I (joint work with Zhiyang Jia) examine the effect of state dependence to policy change responses over time by extending the static discrete choice microsimulation model described in Chapter 2 to an intertemporal setting. The extensions of the model are similar in spirit to contributions by Haan (2010) and Prowse (2012), although the underlying theoretical job choice framework differs slightly. One advantage of the job choice framework is that it allows me to distinguish between state dependence in preferences and state dependence in the set of job opportunities. I estimate the model on Norwegian administrative panel data over the period 2003-2007 for the subset of female wage earners in couples, although the framework can be used in general for all wage earners. I find significant positive state dependence both at the intensive and extensive margin which causes sluggish labor supply responses to various changes in the wage rate and the tax schedule. The simulation results suggest that within the first year of a permanent policy change about 50 percent of the full labor supply response can be expected, whereas 90 percent of the response is reached after about 4 years both at the intensive and extensive margin.

Finally, in Chapter 5, I summarize the main findings. I further discuss possible shortcomings of the applied models and provide some suggestions for further research.

Chapter 2

Validation of a Structural Labor Supply Model by the Elasticity of Taxable Income[1]

2.1 Introduction

Income responsiveness to tax change is a core issue in public economics. This is reflected in numerous estimates based on different methodological approaches. Public finance practitioners are often asked to predict responses to alternative policy changes, and some institutions, such as the Joint Committee on Taxation (U.S.), the Institute for Fiscal Studies (U.K.), and the Research Department of Statistics Norway, are expected to deliver empirical measures of effects to the decision-makers in their respective countries. The application of certain modeling tools is often a prerequisite for this, and the static labor supply model stands out as a practical alternative. Based on cross-sectional observations of households and individuals' consumption and connections to the labor market (typically working hours) it is possible to either directly apply a labor supply function or estimate a utility function; see the review of the literature in Blundell and MaCurdy (1999). The parameter estimates can, in turn, be used to simulate effects on working hours or incomes of hypothetical changes in the tax system.

A second main method of obtaining information about the relationship between income and taxes is based on analyses of observations before and after a realized policy reform. The identification of response estimates typically involves applying the difference-in-differences estimator or related econometric techniques. Treatment effects are measured by exploiting the fact that policy reforms can be seen as defining quasi-experiments. A tax reform generates net-of-tax rate changes along the income scale, often resulting in substantial tax changes for

[1] This chapter is based on joint work with Thor Olav Thoresen, Statistics Norway.

some taxpayers, whereas others are more or less unaffected. The elasticity of taxable income (ETI) is a key concept. It measures the response in taxable income to a change in the net-of-tax rate; see the survey of the ETI literature in Saez et al. (2012).[2]

ETI estimates have limited value in the prediction context, because they often rely on a particular reform for identification. This makes them less applicable when assessing new policy changes.[3] The approach has gained in popularity because of its simplicity: it uses information about incomes, which is normally more easily accessible than data on working hours, and it exploits standard econometric techniques. However, identification is rather challenging which we will return to. The motivation underlying the ETI literature is typically to capture the full set of behavioral responses to taxation, including evasion or avoidance through tax planning. We suggest that the empirical method in the ETI literature also can be used to identify earnings responses that can be compared to the results of micro-simulations based on the structural labor supply model, and can thus represent a source of validation for the labor supply model. Obviously, the ETI methodology relies on certain assumptions and it therefore does not necessarily uncover the true responses to a reform. The aim of the validation exercise is to explore the degree of correspondence between results. It is reassuring if both sources of information indicate similar response magnitudes.

Although the discrete choice labor supply model continues to be a key instrument for predicting policy changes, serious concerns have been raised about the ability of structural models to generate robust predictions about the effects of policy changes; see LaLonde (1986), Imbens (2010), and Angrist and Pischke (2010). As emphasized by Eissa and Hoynes (2006), Todd and Wolpin (2006), Blundell (2006, 2010), Keane (2010b,a), and Heckman (2010), it is essential to use other sources of information to validate the models. For example, Blundell (2006) argues that "...simple difference in difference evaluations can be valuable for validating the specification of more fragile microeconometric models" (p. 425), whereas Keane and Wolpin (2007) argue pragmatically that there is no "true" decision-theoretic model, but only models that are better or worse at addressing particular questions.[4] Moreover, because recent surveys of the static labor supply literature do not seem to agree on how responsive individuals are

[2] By ETI studies, we refer to reduced-form studies developed over the last few decades (after initial contributions by Lindsey, 1987 and Feldstein, 1995). They focus on incomes, in repeated cross-sections or in panels, before and after a major change in the tax schedule, which generates variation in tax changes across individuals.

[3] However, Carroll and Hrung (2005) and Thoresen et al. (2012) are examples of studies that use estimates of the elasticity of taxable income to simulate outcomes.

[4] Brewer et al. (2006), Cai et al. (2008), Hansen and Liu (2011) and Pronzato (2012) are other examples of studies that describe labor supply model validations.

to changes in taxes, even in the static case (see Keane and Rogerson, 2012 and Chetty et al., 2011), it is worthwhile to cross-check the results from two approaches.

If the discrete choice structural labor supply model is able to produce reasonable out-of-sample predictions, we argue that it is thus a powerful tool for improving the informational basis for policy decisions. In this light and given the substantial efforts that have recently been put into obtaining new empirical evidence in both the structural labor supply literature and the ETI literature, we find it surprising that more is not being done to reconcile the evidence and validate results.

Since the main ambition of the present analysis is to show how the results of panel data analysis (using the ETI methodology) can be exploited to validate a structural model, we let earning responses (and not changes in working hours) be the focal point of the discussion. One important message is that validation cannot simply be carried out by comparing average wage elasticities from the labor supply model to average net-of-tax elasticities from the ETI approach; a nonlinear structural labor supply model produces responses that differ along the income scale, which must also be accounted for in a validation exercise.

The structural labor supply model discussed in the present study is a discrete choice model developed by Statistics Norway. It is related to the discrete model of Soest (1995); see Dagsvik and Jia (2012), and Dagsvik et al. (2013). The model is available to Norwegian decision-makers through the model system LOTTE (Aasnes et al., 2007; Thoresen et al., 2010). We focus on validating the model by discussing the response of wage earners at the mid-level and high-end of the income distribution. This particular example is based on the identification of estimates using the ETI approach and exploiting the reductions in top marginal tax rates resulting from the 2006 Norwegian tax reform to derive earned income elasticities. Traditional methods in the ETI literature are used, utilizing the panel structure of data to obtain individual measures of income growth, and employing instrumental variable techniques to obtain measures of change in the net-of-tax rate. To facilitate comparison, we use the discrete choice model to simulate the effects of the specific tax reform on hours of work, and use predicted income levels to obtain a comparable estimate for income elasticity with respect to the net-of-tax rate, which is the key measure in the ETI literature.

This chapter is organized as follows. In Section 2.2, we present the two methodological approaches to obtaining tax response estimates, followed by the presentation of results in Section 2.3. In Section 2.4, we bring the results together and discuss possible caveats. Section 2.5 concludes the chapter.

2.2 Empirical Models for Income and Tax Relationships

A whole range of different response estimates can be found in the labor supply literature, reflecting different theoretical models and methodological approaches. In the present analysis, we discuss evidence from two well-known static approaches:[5] tax simulation based on a structural discrete choice labor supply model, and estimation of the elasticity of taxable income by employing panel data and a quasi-experimental identification strategy. Given that estimation of structural labor supply models often involves severe econometric challenges[6] (see reviews in Blundell and MaCurdy, 1999 and Keane, 2011), reduced-form estimation based on panel data could represent a more straightforward empirical technique for public finance practitioners. However, in addition to the identification methods relying on rather strong assumptions (see for example Moffitt and Wilhelm, 2000), the main limitation of the ETI approach is that the "treatment effect" must be interpreted in terms of the specific tax change under consideration. In general, it is therefore not informative about the effects of other policy changes.

We have recently seen discussions in the literature concerning the advantages of structural modeling versus results derived from quasi-experimental research designs; see, for instance, Chetty (2009b), Angrist and Pischke (2010), Deaton (2010), Heckman (2010), Heckman and Urzua (2010), Imbens (2010), and Keane (2010a,b). As Chetty (2009b) emphasizes, the ETI approach is not easy to place in relation to the two stereotype classifications, since the elasticities it produces share important characteristics with both strands of the literature.[7] For instance, like structural models, the ETI framework departs from an underlying utility-maximizing behavior and produces precise statements about welfare implications. The identification strategy has important similarities with experimental studies, however.[8]

[5] Chetty et al. (2011) refer to this type of evidence as steady-state elasticities. Recent surveys of the literature, such as Blundell and MaCurdy (1999) and Keane (2011), review both results of static approaches and frameworks based on life cycle models.

[6] It can be argued that the discrete choice version of structural modeling is a more practical method than the conventional continuous approach, based on marginal calculus. The structural labor supply model associated with Hausman becomes very complicated when more general and flexible model specifications are used; see Bloemen and Kapteyn (2008).

[7] Chetty (2009b) therefore introduces a third class, the "sufficient statistic" category, which covers studies that make predictions about welfare without estimating or specifying structural models.

[8] The early work of Feldstein (1995) is clearly close to an empirical design that relies on "treatment" and "control" groups, and uses a differences-in-difference estimation technique. However, more recent estimation methods, initiated by Auten and Carroll (1999) and Gruber and Saez (2002), can be seen as standard linear regressions with a first-differenced dependent variable and instrument for the change in the net-of-tax rate. The idiosyncrasy of the results stems from the use of one particular reform to derive estimates, which limits the applicability of estimates for other tax changes. For example, most reforms studied in the literature have involved changes in

Even though there are methodological concerns with both sources of information on tax responses, they enable cross-checking of the empirical results, which, in turn, can be employed to validate the structural model used to simulate the effects of prospective policies. In this section, we present the main characteristics of the two methods of deriving response estimates. First, we present a discrete choice labor supply model. We then describe how tax response estimates can be derived from the analysis of panel data.

2.2.1 The Discrete Choice Model Formulation

Discrete choice models of labor supply based on the random utility modeling approach have gained widespread popularity,[9] mainly because they are much more practical than the conventional continuous approach based on marginal calculus; see Creedy and Kalb (2005) for a survey of the literature and Soest (1995), Bingley and Walker (1997), Blundell et al. (2000), Van Soest et al. (2002), Creedy et al. (2006), Haan and Steiner (2005), Labeaga et al. (2008), and Blundell and Shephard (2012) for applications. The maximization problem for a person in a single-individual household can be seen as choosing between bundles of consumption (C) and leisure (L), subject to a budget constraint, $C = f(hw, I)$, where h is hours of work, w is the wage rate, I is non-labor income, C is (real) disposable income and $f(.)$ is the function that transforms gross income into after-tax household income.

In the empirical specification of the labor supply model applied here, agents are assumed to make choices with respect to "job"; see Aaberge et al. (1995), Dagsvik and Strøm (2006), Dagsvik and Jia (2012) and Dagsvik et al. (2013), where each job is characterized by a discrete set of hours, but several jobs might be characterized by the same working hours. In addition to consumption and leisure, the individual is assumed to have preferences regarding other job characteristics that are unobserved by the researcher. This means that the utility function of the household can be expressed as $U(C, h, z)$, where $z = 1, 2, \dots,$ refers to market opportunities (jobs) and $z = 0$ refers to the non-market alternative. The utility function is assumed to be additively separable, $U(C, h, z) = v(C, h) + \varepsilon(z)$, where $v(.)$ is a positive deterministic function and the random unobserved components $\varepsilon(z)$ are dependent on job z in addition to unobserved

the top marginal tax rate.

[9] Despite its popularity among practitioners of labor supply analysis, less attention is devoted to this framework in recent reviews of the literature. Keane (2011), for example, essentially ignores the (static) discrete choice approach to labor supply altogether. Given that the approach is played down and only referred to as a somewhat crude approximate approach that makes estimation problems manageable, the present analysis, with its emphasis on the "job" notion, holds the promise of a coherent theoretical foundation for the discrete choice labor supply; see Dagsvik and Jia (2012) and Dagsvik et al. (2013) for more details.

individual characteristics. We assume that the random components are i.i.d. extreme value distributed with c.d.f. $exp\,(exp(-x))$ for positive x, which implies independence of irrelevant alternatives (IIA). The strict IIA assumption is weakened, however, by allowing for random effects in relation to the wage rate.[10]

Let $\psi(h) = v\,(f\,(hw,I),h)$ be the representative utility of jobs with hours of work h, a given individual specific wage rate w, and non-labor income I. We further assume that individuals face restrictions on the set of available market opportunities. Let $B(h)$ denote the agent's set of available jobs with hours of work h, and $m(h)$ define the number of jobs in $B(h)$. We assume that there is only one non-market alternative, so that $m(0) = 1$.

Now, let D be the set of possible hours of work. Then, by applying standard results in discrete choice theory (McFadden, 1984), it follows that the probability that the agent will choose job z can be expressed as

$$P\left(v(f(hw,I),h) + \varepsilon(z) = \max_{x\in D} \max_{k\in B(x)} \left(v(f(xw,I),x) + \varepsilon(k)\right)\right) = \frac{exp\psi(h)}{\sum\limits_{x\in D}\sum\limits_{z\in B(x)} exp\psi(x) + exp\psi(0)} \tag{2.1}$$

However, since we are not able to observe the particular job choice, we derive an expression for the probability of choosing any job with hours of work h by adding all the alternatives within $B(h)$.

$$\varphi(h) = \sum_{z\in B(h)} \frac{exp\psi(h)}{\sum\limits_{x\in D}\sum\limits_{z\in B(x)} exp\psi(x) + exp\psi(0)} = \frac{exp\psi(h)m(h)}{exp\psi(0) + \sum\limits_{x\in D} exp\psi(x)m(x)} \tag{2.2}$$

When $h = 0$, we get

$$\varphi(0) = \frac{exp\psi(0)}{exp\psi(0) + \sum\limits_{x\in D} exp\psi(x)m(x)} \tag{2.3}$$

The number of jobs with hours of work h, $m(h)$, can be decomposed into $\theta_i g(h)$, where θ_i defines the total number of jobs available to the individual and $g(h)$ is the fraction of jobs available to the agent with offered hours of work equal to h. We will call $m(h)$ the opportu-

[10] The wage rate is replaced by a wage equation that includes a stochastic error term, and thus a mixed multinomial logit model follows; see McFadden and Train (2000) and Haan (2006).

nity measure and $g(h)$ the opportunity distribution. By inserting the decomposed opportunity measure into the expressions for probabilities, we obtain

$$\varphi(h) = \frac{exp\psi(h)g(h)\theta}{exp\psi(0) + \theta \sum_{x \in D} exp\psi(x)g(x)} \tag{2.4}$$

and

$$\varphi(0) = \frac{exp\psi(0)}{exp\psi(0) + \theta \sum_{x \in D} exp\psi(x)g(x)} \tag{2.5}$$

The resulting expression is a choice model that is analogous to a multinomial logit model with representative utility terms, $\psi(h)$, weighted by the frequencies of available jobs, $m(h) = \theta g(h)$. To identify the model, we assume for the sake of simplicity that the number of jobs available, ϑ_i, is a function of years of education and that the opportunity distribution, $g(h)$, is constant over h apart from a possible peak for full-time. The empirical specification of this model turns out to be similar to van Soest's model (1995), although the rationalization for introducing state-specific dummy variables is an important extension.

Appendix B shows how $\psi(h)$ and the wage rate are specified. It presents the estimation results for single males, single females, and, separately, for males and females in couples (married/cohabiting). They are utilized in the simulation of labor supply responses to the Norwegian tax reform of 2006, presented in Section 2.3.

2.2.2 Utilizing Direct Observations of Income Growth

The approach taken in much of the ETI literature departs from an underlying utility-maximizing behavior similar to that seen in the standard labor supply literature above (Feldstein, 1999; Blomquist and Selin, 2010; Saez et al., 2012). Individuals are assumed to maximize a utility function that increases in consumption (C) and decreases in taxable income (q), subject to a budget constraint described by $C = (1 - \tau)q + R$, where τ is the marginal tax rate (which applies to a linear segment of the tax schedule) and R is virtual income. Accordingly, the "supply function" of taxable income is estimated as a function of the marginal tax rate and virtual income. The formulation thus suggests a closer relationship to the part of the structural labor supply literature that is based on estimation of a continuous labor supply function with a piecewise-linear budget constraint, as in Burtless and Hausman (1978), and Hausman (1985).[11]

Moreover, whereas standard labor supply approaches usually focus on the choice of hours of work (h) given an individual-specific wage rate, a main motivation for the ETI approach is that it allows for a broader range of responses to changes in marginal tax rates, such as tax avoidance and evasion captured by the taxable income response, as denoted by Feldstein (1995). In the present context, we define taxable income q as earned income, defined as the wage rate (w) times hours of work (h).[12] Earnings responses to the marginal tax rate can be identified since we analyze a reform period with changes in the tax schedule for labor income.[13]

Panel data covering a period of net-of-tax rate variation across individuals and across time (often covering a tax reform) have been the main data source for the identification of ETI estimates. We let income for individual i at time t, q_{it}, be explained by a time-specific constant, κ_t, the net-of-tax rate, $log(1-\tau_{it})$, unobserved heterogeneity μ_i and the remaining iid error term, ξ_{it},

$$log\, q_{it} = \kappa + \lambda log(1-\tau_{it}) + \mu_i + \xi_{it} \tag{2.6}$$

The basic framework for identification in the ETI literature consists of various estimations of a first-differenced version of (2.6), using panel data for two periods:

$$\triangle log\, q_{it} = \kappa + \lambda \triangle log(1-\tau_{it}) + \triangle \xi_i \tag{2.7}$$

The coefficient of interest, λ, measures the elasticity of income with respect to changes in the net-of-tax rate defined as $\frac{1-\tau}{q}\frac{dq}{d(1-\tau)}$. The reliability of results depends on carefully framed empirical designs for the identification of the key parameter, including controls for individual characteristics that might affect income growth. One main methodological identification challenge (w.r.t. λ) has been the endogeneity of the tax variable, which has led to the estimation of (2.7) using IV techniques, for instance employing the difference-in-differences estimator, and

[11] The Hausman approach thus deviates from the standard discrete choice model (Soest, 1995), in which estimation is carried out directly on the utility specification.

[12] A separation into measures of wage and working hours is not possible for the panel data source which we use to obtain ETI estimates for earnings.

[13] In the Norwegian context, labor income is not subject to deductions and therefore equals taxable labor income, so that responses in the form of tax avoidance and deduction behavior are not relevant.

grouping individuals into treated and non-treated groups based on pre-reform income levels. Feldstein (1995) is an example of this.[14] Many post-Feldstein studies employ a closely related exclusion restriction, using the change in net-of-tax rates based on a fixed first period income as the instrument in an IV regression; see Auten and Carroll (1999) and Gruber and Saez (2002). Thus, as already noted, the ETI literature is related to methods commonly used in the "experimentalist" or "program evaluation" literature; see, for instance, Imbens and Wooldridge (2009). However, as tax reforms typically involve a reduction or increase in top marginal tax rates and small or no changes at lower income levels, the treatment and control groups follow from their income levels. We are thus far from an ideal randomized trial situation.

The estimated elasticity can be interpreted as the average treatment effect for the treated. In other words, if we let a parameter be a zero-one indication of being treated (experiencing net-of-tax rates changes, or not), as in Feldstein (1995), we identify $E(\lambda|\delta_{it}=1)$. According to Blundell and MaCurdy (1999), this parameter is subject to conventional sample selection biases and cannot as a rule be used to simulate policy responses.[15] Irrespective of this discussion, we focus on the use of ETI as a quasi-experimental method to validate the predictions from a structural model, as suggested by Blundell (2006).

2.3 Tax Response Estimates

In this section, we probe deeper into the cross-checking of the results of the two methodologies, discuss the empirical content of the two sources of information and, finally, assess the validity of the structural discrete choice model. The change in marginal tax rates for wage income as a result of the Norwegian tax reform of 2006 is used to illustrate the effects. After presenting the tax reform, which serves as a tool for the identification of tax behavioral responses, we present the evidence from the panel data approach, and these results are then compared to the predictions of the labor supply model.

[14] Feldstein (1995) used a table version of this technique. Aarbu and Thoresen (2001) employed a regression version of the same procedure as one of two econometric methods.

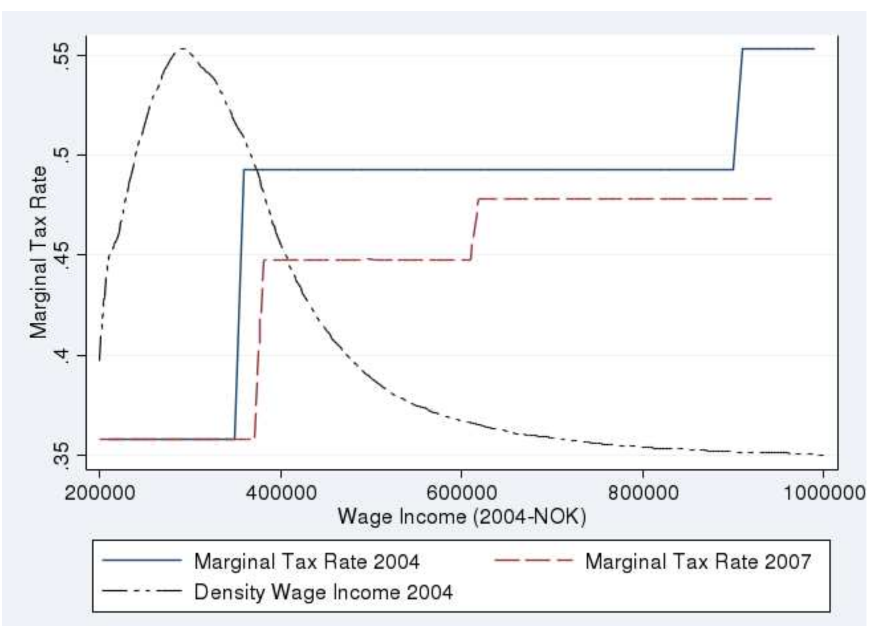
[15] The estimated elasticities can only be used to simulate hypothetical tax reforms under the assumption that the elasticity is constant over the income distribution, which is clearly not consistent with findings from the structural labor supply literature.

2.3.1 The Norwegian Tax Reform of 2006

Norway has a “dual income tax” system, enacted through the 1992 tax reform, which consists of a combination of a low proportional tax rate on capital income and progressive tax rates on labor income. The system proliferated in the Nordic countries in the early 1990s. The Norwegian version had a flat 28 percent tax rate levied on corporate income, capital income and labor income coupled with a progressive surtax applicable to labor income. The gap between marginal tax rates on capital income and wage income was problematic, and the schedule was reformed in 2006 in order to narrow the differences by introducing a shareholder income tax, and, most importantly in the present context, by cutting marginal tax rates for labor income.

The tax reform was gradually implemented in 2005 and 2006; in Figure 2.1 we compare schedules for 2004 (pre-reform) and 2007 (post-reform). The figure shows the principal features of the Norwegian labor income tax system: a two-tier surtax that supplements a basic income tax rate of 28 percent plus a 7.8 percent social insurance contribution. In 2004, the first tier of the surtax was applied to incomes above NOK 354,300 at a rate of 13.5 percent, and the second tier of 19.5 percent applied to income in excess of NOK 906,900. The reform meant that the maximum marginal tax rate fell from 55.3 to 47.8 percent, but became effective at a lower level.

Figure 2.1: Reductions in marginal tax rates as a result of the tax reform



2.3.2 Evidence from Panel Data Estimations

In the following, we closely follow the conventional approach in the ETI literature; see, e.g., Gruber and Saez (2002). The main data source is the Income Statistics for Persons and Families (Statistics Norway, 2005), a register-based data set that covers the whole Norwegian population, with data from income tax returns as the main component. We limit the data set to wage earners over the period 2000-2008,[16] utilizing overlapping three-year individual differences. More details on the empirical specification and sample restrictions are provided in Appendix A.

Changes in net-of-tax rates are instrumented by using the tax change for a constant base year income (referring to the first year in each three-year difference). Since the tax instrument is constructed from the base year income and the dependent variable is growth in income (over the same three-year period), a control is necessary for mean reversion and drifts in the income distribution. Mean reversion in this context refers to an observed negative correlation between initial income and income growth due to transitory shocks in income, whereas changes in the income distribution can lead to correlations in both directions. These phenomena should not be mistakenly attributed to the exogenous tax change. As a solution, Auten and Carroll (1999) included the base year income in logs as an additional explanatory variable, whereas Gruber and Saez (2002) extended this approach by allowing for a piecewise linear function of base year income. Here, 10 linear splines or a third degree polynomial of base year income are used to control for mean reversion and drifts in the income distribution.[17]

Table 2.1 shows the results of the 2SLS regressions. To address the mean reversion problem, in the first two columns, we have included 10 splines of log income, and, in the third and fourth column, a third degree polynomial of log income. Specifications (2) and (4) include a control for virtual income, following Blomquist and Selin (2010); see also Appendix A. Although results (in general) are sensitive to the inclusion of the mean reversion control, there is only a small difference between the estimates including 10 splines or a third degree polynomial of base year income. The uncompensated elasticity of earnings with respect to net-of-tax is estimated to be about 0.05–0.06 without the income effect, and 0.03–0.04 after the income effect is controlled for. The income elasticity is small and negative, as expected.

[16] We analyze a somewhat longer time period than only the actual tax reform period (2004-2007) in order to improve the estimates for the non-tax-related control variables.

[17] In addition we alleviate the problem of mean reversion by excluding observations with low base year income (below percentile 33); see Appendix A.

Table 2.1: Estimates of the net-of-tax rate elasticity for earned income. 2SLS regression results for all wage earners

	Mean reversion control			
	Splines		Polynomial	
	(1)	(2)	(3)	(4)
Net-of-tax rate elasticity	0.0562***	0.0370***	0.0531***	0.0356***
	(0.0023)	(0.0032)	(0.0023)	(0.0031)
Income elasticity		-0.0091***		-0.0105***
		(0.0012)		(0.0012)
Number of observations	4,933,291	4,331,276	4,933,291	4,331,276

Note: Standard errors in parentheses. All regressions include control variables for gender, wealth, age, age squared, marital status, number of children under and above the age of 6, newborn children, residence in Oslo/ densely populated area, non-western origin, years of education, 9 dummies for field of education, income shifting control and year dummies. Full regression output is reported in Table 2.A1, Appendix A.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The estimated (average) net-of-tax elasticities are small compared to most other ETI studies. According to Saez et al. (2012), estimates from the U.S. (after Feldstein, 1995) range from 0.12 to 0.40. Our estimates, however, measure the responses in wage earnings only, and do not identify responses in avoidance and reporting behavior. The estimates are in line with Kleven and Schultz (2011), who report elasticities of approximately 0.05 for wage earners in Denmark.

Further, we divide the sample into four groups: single females, single males, females in couples, and males in couples. Response estimates for specific groups facilitate closer comparison with the simulation results from the structural model estimation. A third degree polynomial is used as a mean reversion control and the additional controls for virtual income[18] and income shifting (see Appendix A) are excluded. The results in Table 2.2 suggest that the responses are small positive, 0.04-0.05, but statistically significant for all four groups of wage earners.

[18] Income effects are often neglected in the ETI literature, under the assumption that the effect is close to zero, e.g. as found in Gruber and Saez (2002). Moreover, there is no standardized method of constructing income controls. We have relied on a method proposed by Blomquist and Selin (2010) that includes non-labor income, and therefore seems most appropriate for our setting. However, it turns out to be problematic that the two excluded instruments for net-of-tax rate and virtual income are similarly constructed and therefore appear to suffer from a problem of collinearity, in particular when categorizing into homogenous groups of individuals. We have therefore omitted income effects in Table 2.2.

Table 2.2: Estimates of the net-of-tax rate elasticity for earned income. 2SLS regression results for separate groups of wage earners

	Single females	Single males	Females, couple	Males, couple
Net-of-tax rate elasticity	0.0377***	0.0395***	0.0441***	0.0547***
	(0.0061)	(0.0059)	(0.0049)	(0.0031)
Number of observations	576,232	959,151	1,109,651	2,287,960

Note: Standard errors in parentheses. All regressions include control variables for gender, wealth, age, age squared, marital status, number of children under and above the age of 6, newborn children, residence in Oslo/densely populated area, non-western origin, years of education, 9 dummies for field of education, income shifting control and year dummies. Full regression output is reported in Table 2.A2, Appendix A.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

2.3.3 Microsimulation Results from the Discrete Choice Model

Next, we show how we can derive estimates of comparable net-of-tax rate elasticities from a labor supply model simulation to assess to which extent the discrete choice model replicates the results of the ETI analysis.

The discrete choice structural model is estimated by using information on hours of work from the Labor Force Survey (Statistics Norway, 2003) and income data from the Income Statistics for Persons and Families (Statistics Norway, 2005) for 2004 (a pre-reform year). Four separate specifications, for men in couples, women in couples, single women, and single men, are estimated. The results of the labor supply model estimations are presented in Appendix B, including the results of the estimations of wage rate equations.

Before addressing the results of simulations of the earned income elasticity with respect to the net-of-tax rate, we present standard wage elasticities from the estimated model. The uncompensated wage elasticities are obtained by increasing the gross hourly wage by one percent and using the model and parameter estimates to simulate the percentage change in predicted hours worked for each individual. The average elasticity for each group is shown in Table 2.3. Further, the wage elasticity is decomposed into a participation elasticity and an elasticity conditional on participation, measuring the extensive and intensive margin, respectively. The results for the intensive margin are most relevant with respect to the results of the ETI framework. They show modest elasticities, in the range 0.10–0.29.

Note that the elasticity of working hours with respect to the wage rate (for a given tax rate)[19]

[19] To be accurate, the wage elasticities measured above are gross wage elasticities, where the increase in wages might increase the tax rate as well.

Table 2.3: Wage elasticity estimates derived from simulation of structural labor supply model

		Wage elasticities	
	Total	Extensive margin	Intensive margin
Males in couples	0.16*** (0.0081)	0.005*** (0.0013)	0.15*** (0.0079)
Single males	0.11*** (0.0137)	0.010*** (0.0019)	0.10*** (0.0121)
Females in couples	0.38*** (0.0107)	0.090*** (0.0051)	0.29*** (0.0079)
Single females	0.21*** (0.0197)	0.052*** (0.0076)	0.16*** (0.0195)

Note: Standard errors in parentheses obtained by non-parametric bootstrapping, 30 repetitions.

is conceptually equal to the elasticity of earnings with respect to the net-of-tax rate (for a given wage rate). However, as the labor supply model is non-linear, obtaining net-of-tax rate elasticities (as measured by the ETI method) requires that the particular reform used for identification must be taken into account. To approach comparable measures, we therefore let the results of labor supply model simulations enter into a regression, similar to that seen in the ETI literature (see Section 2.2.2). First, the structural model is used to simulate the pre-reform and post-reform working hours; average estimates for the four groups of wage earners are shown in Table 2.4. Growth in labor income is identical to growth in predicted hours for an individual-specific wage rate. Then, measures of the change in the net-of-tax are derived from the predicted income levels, and instrumented using similar methods as in the ETI literature by using the change in net-of-tax for constant (predicted) initial labor income (predicted pre-reform hours times the individual's constant wage rate);[20] see the results in Table 2.5.

Table 2.4: Predicted weekly hours of work, pre- and post-reform, derived from simulation of labor supply model

	Pre-reform working hours	Post-reform working hours	Difference
Males in couples	38.96 (0.041)	39.27 (0.051)	0.81% ***
Single males	39.01 (0.078)	39.25 (0.088)	0.62% *
Females in couples	36.26 (0.066)	36.39 (0.069)	0.41%
Single females	37.22 (0.088)	37.33 (0.093)	0.30%

Note: Standard errors in parentheses obtained by non-parametric bootstrapping, 30 repetitions.

[20] Predicted working hours follow from the individual's probability distribution, by a draw from a uniform distribution (the same draw applies for each individual pre- and post-reform). An alternative method is to use the expected predicted working hours estimate for each individual pre- and post-reform. This leads to similar results, although the income distribution becomes more compressed. As in the panel data analysis, the regression is restricted to individuals with predicted pre-reform income in percentile 33 or above.

Table 2.5: Estimates of the net-of-tax rate elasticity derived from labor supply model simulation

	Net-of-tax rate elasticity	Std. error
Males in couples	0.092^{***}	(0.0051)
Single males	0.076^{***}	(0.0068)
Females in couples	0.055^{***}	(0.0037)
Single females	0.052^{***}	(0.0039)

Note: Standard errors obtained by non-parametric bootstrapping, 30 repetitions.

In the next section, we present a comparison of the results of Table 2.5 with the results of the panel data analysis. At this stage, we observe that the net-of-tax rate elasticities are lower than the wage elasticities (see Table 2.3), ranging between 0.05 and 0.09. Although the estimated wage elasticities were clearly higher for women in couples (0.29 versus 0.10–0.15 for males), the estimates of the net-of-tax rate elasticities suggest that responses are larger for males than females. This indicates that males at the high end of the income distribution are on average more sensitive than high-income females to the particular change in the tax schedule exploited to derive the net-of-tax rate elasticities (changes in the top marginal tax rates).

2.4 Reconciling the Evidence

As described in the previous section, we obtained net-of-tax rate elasticities from the discrete choice model that can be compared to the results of a traditional ETI analysis. One should be aware, however, of important differences between the discrete choice labor supply model and the underlying framework leading to the ETI approach. Before comparing the results of the two approaches, let us therefore review some of the main differences, such as discrete/continuous choice, responses through working hours/taxable income, the underlying time frame and, more generally, the distinction between a structural approach (used for simulation) and a reduced form panel data analysis.

Firstly, the structural labor supply model we have estimated is based on discrete choice instead of marginal optimization, generating a probability distribution for different working hours options.[21] There are different procedures that can be employed in the simulation of such models,

[21] The probability distribution follows from McFadden's conditional logit framework; see for instance McFadden (1984). Recall that the particular labor supply model presented here is a "job choice" model, which is turned into a choice between different categories of hours of work.

all of which respect the probabilistic nature of the model. The present model describes the effects of alternative policies by letting the overall probability distribution be altered as the economic conditions change. Since it is the probability distribution that describes the individual's working hour choices of different policy alternatives, all individuals are affected by a reform to some extent. In the ETI literature, the response estimates reflect (in a somewhat simplified manner) the policy change exploited to obtain estimates, dividing the data into "treated" and "less or not treated", based on marginal optimization. In this perspective the modeling of the ETI literature is therefore more related to the perspective of continuous hours structural labor supply models such as the so-called Hausman model; see Section 2.2.2.

Secondly, the models differ in the type of responses that tax changes induce. As already emphasized, the ETI literature includes a whole range of responses, including tax planning and tax avoidance, as it typically focuses on total taxable income. In our study, we have a more narrow focus on wage earners' responses in the form of labor income (hourly wage times hours). We should still capture responses in both working hours and the wage rate. In the ETI literature, the wage rate is seen as a choice variable for the individual, as he or she can alter his/her wages through increased efforts per hour or by changing jobs. In the labor supply literature, the wage rate is typically considered to be fixed at the individual level. There have been some attempts to endogenize the wage rate in labor supply models (see for example Moffitt, 1984), but they still do not assume that the wage rate can be altered by decisions relating to individual effort.

Thirdly, the methods differ with respect to the time frame. The structural model is a static model where a new long-run steady-state is immediately attained. To obtain the estimates of the ETI, on the other hand, we use the ad hoc choice of three-year spans. Because the structural model might be inappropriate for describing short-term responses (see Chapter 4), it is not obvious how such results can be compared to the standard time framing in the ETI literature.

Finally, it is important to warn against giving precedence to either of the two empirical approaches presented here. The structural labor supply model is based on a model for optimizing behavior, whereas the panel data analysis yields an average effect for the treated individuals who were subject to the particular reform used for identification. The advantage of the structural approach is that the model can be used for any hypothetical tax reform, and it should be valid for any time period because it endeavors to estimate the deep underlying structural parameters. However, because the model might be too simple or suffer from misspecification, the data-driven results of the ETI approach might serve as a test of how well the structural model performs. As is evident from the experience of ETI estimations, however, the results cannot be characterized as uncovering "true responses". Ideally, we would not only require pre- and post-reform data, but also counterfactual income levels in the case where no reform occurred. Given the lack of counterfactuals, one of the main practical problems of the ETI approach we have adopted here is that the tax rate instrument is correlated with other explanatory variables for wage growth, such as mean reversion and trends in the income distribution, that are unrelated to the tax reforms.

Despite the major differences in the methodological framework, ETI estimates represent an information source for validation of the simulation model, and in Table 2.6 we restate the comparable results of the structural model and the experimental panel data estimation.[22] The measured net-of-tax rate elasticities are small in both the structural and the panel data analysis, within the range 0.04–0.09. The results are in line with the argument that high-income individuals are less responsive to tax changes, as there is a natural or institutional limit on working hours per week. Our estimates are smaller than typically found in the ETI literature, possibly because we focus on wage earnings in contrast to overall taxable income. Moreover, we look at a strictly defined group of prime age wage earners, with wage income in the median and upper part of the income distribution. This group might be less responsive than self-employed people, people with capital income, and individuals with a weaker attachment to the labor market. In general, it might be argued that the Norwegian institutional setting produces smaller elasticities. The argument presented in Slemrod and Kopczuk (2002) can be used in support of the notion that Norwegians are less responsive.[23]

It is surprising that the structural model predicts somewhat larger responses than the panel data approach; we would expect the panel data estimates to be larger, since a broader measure of responses is arguably captured through earnings. However, the magnitudes of the estimates are indeed very similar. Moreover, larger estimated responses for married males than for females are in line with the predictions of the labor supply model for the particular tax reform under consideration.

[22] For both methods, we estimate the uncompensated elasticities. The income effect is typically estimated to be small in the ETI literature, and it is often assumed that the compensated and uncompensated elasticities are similar; see e.g. Saez et al. (2012). Measures of compensated elasticities are rare in the discrete choice structural labor supply literature; see, however, Dagsvik and Karlström (2005) for a method of obtaining compensated effects.

[23] For instance, in order to be able to uphold a progressive tax system, egalitarian societies may establish institutions to reduce tax avoidance.

Table 2.6: Comparison of net-of-tax rate elasticity estimates from structural labor supply model simulation and analysis of panel data

	Structural model		Panel data	
Males in couples	0.092***	(0.0051)	0.055***	(0.0031)
Single males	0.076***	(0.0068)	0.040***	(0.0059)
Females in couples	0.055***	(0.0037)	0.044***	(0.0049)
Single females	0.052***	(0.0039)	0.038***	(0.0061)

Note: Standard errors in parentheses.

2.5 Conclusion

The discrete choice labor supply model is a tool that is frequently used to analyze a wide range of hypothetical tax and benefit reforms. Given its key role in the decision-making process, it is important to validate its capacity to provide reasonable descriptions of the effects of prospective policies. There has recently been growing interest in validating discrete choice structural models using natural experiments. However, we have yet to see any detailed discussion of how the standard structural labor supply model can be validated using results from the ETI literature. From this perspective, the present study offers a procedure for comparison.

A validation that is simply based on comparisons of average wage elasticities from the labor supply model with average net-of-tax rates from the ETI approach is in danger of being misleading. The reason is that ETI estimates are derived from specific tax reforms, and that they therefore measure the average effects for the individuals treated by the reforms. The non-linearity of the labor supply model, on the other hand, implies different responses along the income distribution.

In this study, we have shown how a version of the labor supply model made available to Norwegian decision-makers (through the model system LOTTE) is validated by ETI estimates. The model is used to simulate the labor supply effects of the Norwegian tax reform of 2006. Earnings are simulated pre- and post-reform under an exogenous wage assumption, and the regression framework of the ETI literature is used to obtain a net-of-tax rate elasticity for the simulated earnings level. These estimates have then been compared with ETI estimates obtained in the conventional manner, using panel data of actual labor income levels before and after the reform.

Our main finding is that simulations from the structural labor supply model yield net-of-tax elasticity estimates that are close to the elasticity estimated on the basis of the panel data, ranging from 0.05 to 0.09 for the structural model and 0.04 to 0.055 for the panel data analysis. We thus find it reassuring that the predictions of the labor supply model are not far from the results of the panel data analysis.

Even though some doubts have been expressed about the capacity of structural models to predict outcomes of policy changes, it continues to be a main tool for public finance practitioners. Instead of dismissing the approach as a means of obtaining policy guidance, more effort should be put into qualifying models through validation. The present study is just one example among many possible certifications of this key instrument.

Appendix A. Empirical Specification for Panel Data Analysis

As discussed in Section 2.2.2, the standard framework for estimation of the elasticity of taxable income is to employ panel data information, estimating a model in differences, typically for a three-year span. We stack observations for each three-year difference (2000-2003, 2001-2004, ... , 2005-2008) over the period 2000-2008, and add time invariant explanatory variables as possible explanations for income growth. Thus, the three-year difference in (log) taxable income, q_{it}, is explained by a period-specific effect, κ_t, differences in marginal tax rates, $1-\tau_{it}$, and a set of control variables, X_{it}.

$$log\left(\frac{q_{it+3}}{q_{it}}\right) = \kappa_t + \lambda_1 log\left(\frac{1-\tau_{it+3}}{1-\tau_{it}}\right) + X_{it}\omega + \xi_{it} \tag{2.A1}$$

The key parameter is λ_1, which measures the uncompensated elasticity of taxable income.

We estimate the model by using panel data derived from administrative registers, with Income Statistics for Persons and Families as our main source (Statistics Norway, 2005). The income register contains information for the whole population of Norway (about 4.6 million in 2004). As for the structural model, we limit the sample to wage earners aged 25-62 years, defined as having wage income as their main source of income, and we exclude individuals with positive income from self-employment or pensions. In addition, we limit the sample to individuals with taxable labor income above percentile 33 (about NOK 250,000 in 2004) in the base year (the first year in each three-year difference). We are left with about five million three-year differences.

There are two reasons for excluding the lower income levels. Firstly, we are mainly interested in the effect of decreased surtax rates, which only affect about 1/3 of the wage earners. Secondly, the mean reversion problem (described in more detail below) is especially severe for individuals who initially had low income, which makes this group less appropriate as a control group.

The dependent variable is growth in gross labor income, the same income concept that forms the tax base for labor taxation. The elasticity we obtain can therefore be denoted as the elasticity of earned income. The actual marginal tax rate is not immediately available in the data set, but is constructed by tax simulation based on the additional taxes levied on the individual if the income is increased by five percent. The change in marginal tax rate is clearly endogenous, since the marginal tax rate (as a function of income) is jointly determined with income. The tax rate change, $log\left(1-\tau_{it+3}(q_{it+3})/1-\tau_{it}(q_{it})\right)$, is therefore instrumented by a tax rate change for a "constant" or inflation-adjusted initial income level, $log\left(1-\tau_{it+3}((1+b)q_{it})/1-\tau_{it}(q_{it})\right)$, where b corresponds to median income growth over the period t to $t+3$.

The error term of equation (2.A1) is most likely correlated with first period income, q_{it}, for instance because of mean reversion and drifts in the income distribution (Moffitt and Wilhelm, 2000). For example, some individuals with high income in period t and therefore (mistakenly) placed in the treatment group with large reductions in marginal tax rates, will return to their normal income level in period $t+3$, and an income reduction will be recorded. To account for the mean reversion bias, Auten and Carroll (1999) suggest adding $log\, q_{it}$ as an additional control variable. As shown in many analyses, Aarbu and Thoresen (2001) included, this control has a big influence on tax elasticity estimates, and it shifts estimates of the change in the net-of-tax rate from negative to positive. Gruber and Saez (2002) suggest extending the base period income control technique by including a piecewise linear function of $log\, q_{it}$.

The main problem of employing rich controls for mean reversion based on first-period information is that identification of the effect of the net-of-tax rate may become blurred, because the mean reversion control and the tax change instrument depend on the same variable; see, for instance, Saez et al. (2012). The problem is alleviated by including periods both with and without tax changes. Our empirical study also benefits from having other sources of variation in the tax rate than income alone, in view of two tax classes and a separate schedule for northern Norway.

The spline or polynomial function in the log of first period income is not just a control for mean reversion effects along the income scale; it can also be seen as accounting for changes in the income distribution. For example, a trend towards increasing inequality may result in a spurious correlation between lowered tax rates for high-income individuals and income growth rates.

Individual characteristics are included to control for non-tax-related income evolution over time or over the life cycle. We have had access to a number of socio-demographic characteristics, such as age, years of education, field of education, marital status, number of children, geographical location, and area of origin. The regression output confirms that the presence of young children seems to limit income growth, whereas length of education has a positive effect, presumably due to a steeper increase in earnings over the life cycle.

The Norwegian tax reform of 2006 reduced the tax advantages enjoyed by capital income compared to labor income, and it could therefore result in a reversed income-shifting effect where individuals again increase their labor earnings at the expense of capital income; see Thoresen and Alstadsæter (2010) for the measurement of the opposite effect. In the pooled regressions, we control for this possible effect by including an interaction of the individual's log capital income in the base year period with the flat tax rate change in capital income over the three-year period under consideration.[24]

The income effects are often neglected in the literature, since the income elasticity is assumed to be small (as shown by Gruber and Saez, 2002). To test this, we include a virtual income control in some of the main specifications. Like Blomquist and Selin (2010), we construct virtual incomes using procedures seen in the labor supply literature, based on piecewise linear approximations to the budget constraint (see Burtless and Hausman, 1978). Virtual income, $R_{it} = I_{it} + (\tau_{it} q_{it} - \nu_{it}(q_{it}))$, is expressed as the difference between paying the marginal tax on overall labor income, $\tau_{it} q_{it}$, and the actual taxes paid, expressed as $\nu_{it}(q_{it})$. This difference is positive in a progressive tax system with tax allowances. In addition, since q_{it} only captures labor income, non-labor income, I_{it}, is included and assumed to be exogenously given. In non-labor income, we include transfers such as child benefit and other social transfers, in addition to net-of-tax capital income. For couples, non-labor income includes the disposable income of the spouse. Since the model is estimated in first differences, the change in virtual and non-labor income, R_{it}, is instrumented by (again) using the exogenous tax rate change for a fixed income level.

Table 2.A1 shows the full regression output of our main results. In the first set of regressions, (1) and (2), we have included log base year income as a linear function, in the second, (3) and (4), as a 10-piece spline, and, in the third, (5) and (6), as a centered third degree polynomial. We present the results both with and without a control for virtual income.

We find specifications (3)–(6) most convincing, as we believe it is not sufficient to include a linear control for the occurrence of mean reversion.[25] We see that the results are less influenced if either splines or polynomials are used as control variables.

Table 2.A2 shows the full regression output when categorizing the sample with respect to gender and cohabitation status. In all regressions, we have included a centered third degree polynomial of the base year income as a mean reversion control. The results show relatively little variation with respect to these categorizations.

[24] Note that this control variable is not endogenous, since the tax rate is flat and therefore identical for all wage earners. The control variable has the value 0 in periods where no capital tax changes occurred.

[25] When only a linear control is included for mean reversion, the estimated elasticity becomes dependent on sample restrictions.

Further, we present robustness checks regarding sample restrictions and choice of time span, since the decisions to include individuals above percentile 33 and to use three-year differences are both questionable. In Table 2.A3, we present results for the net-of-tax rate elasticity for alternative cut-off rules.[26] In the first regression, we include all individuals in percentile 25 or above, and in the third regression we include all individuals in percentile 40 or above. We expect the estimate of the net-of-tax rate elasticity to be independent of this choice, since, irrespective of the cut-off point, individuals in the control group were not affected by the reform. The results uncover that there are very small differences in the estimated net-of-tax rate elasticities with respect to sample restrictions.

The three-year span has been proposed in the literature to allow some time for individuals to respond to tax changes. The choice is ad hoc, however, and here we present the results for alternative spans, one to four years. Again, the third degree polynomial is used as the mean reversion control. The results are relatively robust to alternative spans, with the lowest elasticity of 0.032 for one year differences. The likely reason is that wage earners do not respond immediately to tax changes; see Chapter 3.

Moreover, we have assessed the extent to which the comparison of results from the structural labor supply model and the ETI panel data approach are influenced by sample differences. The estimation of the labor supply model is based on survey information since it requires detailed information on working hours (see Appendix B), whereas the earnings elasticities presented here are derived from a larger panel data set consisting of the complete set of wage earners. We have therefore looked at the representativity of the observations from the Labor Force Survey (which can be identified through common identification numbers in the panel data set) by limiting the sample to these individuals in an ETI panel estimation. We find that the results for this much smaller sample are not statistically different from the results reported above. However, the estimates are less precise due to the considerable smaller sample size (only about one percent of the wage earner population). The results are available upon request.

[26] Tables 2.A3 and 2.A4 are based on the same specification as Table 2.A1 (Specification 5), with third degree polynomials and without virtual income control.

Table 2.A1: 2SLS regression results for all wage earners

	Mean Reversion Control					
	Linear		Splines		Polynomial	
	(1)	(2)	(3)	(4)	(5)	(6)
Net-of-tax rate elasticity	0.0312***	0.0154***	0.0562***	0.0370***	0.0531***	0.0356***
	(0.0021)	(0.0030)	(0.0023)	(0.0032)	(0.0023)	(0.0031)
Virtual income elasticity		-0.0094***		-0.0091***		-0.0105***
		(0.0012)		(0.0012)		(0.0012)
Income shifting control	0.0112***	0.0107***	0.0111***	0.0106***	0.0111***	0.0105***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Male	0.0412***	0.0333***	0.0418***	0.0338***	0.0416***	0.0337***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Wealth	-0.0003***	-0.0002***	-0.0003***	-0.0002***	-0.0003***	-0.0002***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Age	0.0015***	0.0010***	0.0015***	0.0010***	0.0015***	0.0010***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Age squared	-0.0000***	-0.0000***	-0.0000***	-0.0000***	-0.0000***	-0.0000***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Married	0.0114***	0.0097***	0.0113***	0.0095***	0.0112***	0.0096***
	(0.0002)	(0.0003)	(0.0002)	(0.0003)	(0.0002)	(0.0003)
No. newborn children	-0.0592***	-0.0500***	-0.0596***	-0.0503***	-0.0595***	-0.0502***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
No. children under age 6	0.0202***	0.0181***	0.0202***	0.0181***	0.0202***	0.0182***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
No. children above age 6	0.0081***	0.0085***	0.0082***	0.0086***	0.0081***	0.0086***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Non-western origin	-0.0431***	-0.0420***	-0.0432***	-0.0421***	-0.0431***	-0.0420***
	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0007)	(0.0007)
Residence in Oslo	0.0024***	0.0012***	0.0023***	0.0012***	0.0024***	0.0013***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Densely populated area	0.0096***	0.0093***	0.0096***	0.0093***	0.0096***	0.0093***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Years of education	0.0133***	0.0127***	0.0133***	0.0126***	0.0133***	0.0126***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Field of education	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-0.0192***	0.0055	-0.0174***	0.0069	-0.0182***	0.0074*
	(0.0029)	(0.0031)	(0.0034)	(0.0036)	(0.0029)	(0.0031)
Number of observations [a]	4,933,291	4,331,276	4,933,291	4,331,276	4,933,291	4,331,276

a The number of observations is slightly smaller when allowing for virtual income effects. This is mainly because we have conditioned on individuals' cohabitation status being unchanged over the period, as spouse's income is included in virtual income.

Note: Standard errors in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.A2: 2SLS regression results by gender and cohabitation status

	Female, single	Male, single	Female, couple	Male, couple
Net-of-tax rate elasticity	0.0377***	0.0395***	0.0441***	0.0547***
	(0.0061)	(0.0059)	(0.0049)	(0.0031)
Wealth	-0.0002***	0.0001*	-0.0007***	0.0002***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Age	0.0001	-0.0030***	0.0072***	-0.0016***
	(0.0003)	(0.0003)	(0.0003)	(0.0002)
Age squared	-0.0000	0.0000***	-0.0000***	0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Married	0.0175***	0.0067***	0.0137***	0.0132***
	(0.0011)	(0.0010)	(0.0005)	(0.0004)
No. newborn children	-0.1401***	-0.0072	-0.1430***	-0.0134***
	(0.0026)	(0.0040)	(0.00048)	(0.0005)
No. children under age 6	0.0386***	-0.0022	0.0374***	0.0063***
	(0.0018)	(0.0026)	(0.0006)	(0.0003)
No. children above age 6	0.0150***	0.0126***	0.0094***	0.0037***
	(0.0005)	(0.0008)	(0.0003)	(0.0002)
Non-western origin	-0.0329***	-0.0582***	-0.0273***	-0.0518***
	(0.0019)	(0.0020)	(0.0014)	(0.0009)
Residence in Oslo	0.0077***	-0.0070***	0.0109***	0.0009***
	(0.0006)	(0.0006)	(0.0005)	(0.0003)
Densely populated area	0.0121***	0.0058***	0.0120***	0.0087***
	(0.0009)	(0.0007)	(0.0006)	(0.0004)
Years of education	0.0141***	0.0162***	0.0157***	0.0123***
	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Field of education	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
Constant	-0.0371***	0.0434***	-0.2217***	0.1141***
	(0.0074)	(0.0071)	(0.0063)	(0.0042)
Number of observations	576,232	959,151	1,109,651	2,287,960

Note: Standard errors in parentheses. Third degree polynomial of base year income is used as mean reversion control.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.A3: Robustness checks: Sample restrictions

	Above percentile 25	Above percentile 33	Above percentile 40
Net-of-tax elasticity	0.0520***	0.0531***	0.0534***
	(0.0023)	(0.0023)	(0.0022)
Number of observations	5,486,168	4,933,291	4,439,785

Note: Standard errors in parentheses. All regressions include control variables for gender, wealth, age, age squared, marital status, number of children under and above the age of 6, newborn children, residence in Oslo/ densely populated area, non-western origin, years of education, dummies for field of education, income shifting control, year dummies and third degree polynomial of base year income. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.A4: Robustness checks: Time-span

	One year	Two years	Three years	Four years
Net-of-tax elasticity	0.0320***	0.0418***	0.0531***	0.0463***
	(0.0023)	(0.0022)	(0.0023)	(0.0026)
Number of observations	7,375,466	6,080,466	4,933,291	3,960,093

Note: Standard errors in parentheses. All regressions include control variables for gender, wealth, age, age squared, marital status, number of children under and above the age of 6, newborn children, residence in Oslo/ densely populated area, non-western origin, years of education, dummies for field of education, income shifting control, year dummies and third degree polynomial of base year income. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix B. Specification and Estimation of the Discrete Choice Model

The discrete choice model presented in Section 2.2.1 is estimated for single females, single males and for coupled females and males. For persons in couples, we also estimate individual models, but take the income of the spouse into account by including their income as non-labor income.

To simplify the choice, we group jobs into 11 categories based on weekly hours of work:

$$h_w \in [0-5, 5-10, 10-15, \ldots, 45-50, 50+]$$

As noted in Section 2.2.1, the particular job choice model involves incorporating differences in opportunities into the labor supply modeling, represented by individual differences, θ_i, and variations in opportunities with respect to working hours, $g(h)$. We assume that the densities of offered hours are uniform except for peaks at full-time (35-40), whereas we let the individual differences be determined by years of education (S).[27]

The deterministic part of preferences is represented by the following "Box-Cox" type utility function,

$$v(C,h) = \alpha_0 \frac{(C-C_0)^{\alpha_1} - 1}{\alpha_1} + (\beta_0 + \gamma X) \frac{(\bar{h}-h)^{\beta_1} - 1}{\beta_1}$$

where C measures the household-adjusted consumption level, constructed by dividing the couple or individual's disposable income by $\sqrt{N}$, where N is the number of individuals in the household (including children under 18). C_0 represents the minimum or subsistence household-adjusted consumption level, here set to NOK 60,000. $\bar{h}$ is defined as 80 hours per week and h is working hours per week, so that $(\bar{h}-h)$ measures leisure time. X is a vector of taste-modifying variables including age, education level, children, residence, and origin.

Information about actual and formal working hours in primary and possible secondary jobs and information about labor market status are obtained from the Labor Force Survey of 2004 (Statistics Norway, 2003). This is the main source of labor market statistics in Norway, providing information about approximately 24,000 individuals. Each respondent is asked about hours of work and attachment to the labor market in a reference week over eight consecutive

[27] θ_i is not identified for males due to the low frequency of non-participation.

quarters. Information about incomes, family composition, number of children, education, etc. is obtained from the Income Statistics for Persons and Families (Statistics Norway, 2005) and merged with the Labor Force Survey, using unique personal identification numbers. Based on information about labor force status, we have included wage earners and "potential" wage earners, coded as employed and home workers. Unemployed, self-employed, disabled persons and students are excluded from the sample. We further limit the sample to persons aged between 25 and 62 years, and we define a person as non-participating if he or she works less than five hours per week.

Working hours are measured as actual hours of work in both the primary and secondary job, using the average of the reference week information for four quarters. It is a key assumption that this average is a good proxy of a "normal" working week during the year. An alternative to this measure of working hours is to use contractual hours of work, but we then loose some of the variation in working hours and introduce a possible measurement error in the calculation of the wage rate. The reason is that individuals who normally work overtime, might be paid for that through their standard wage or have the option of charging employers for their extra workload. We see it as important to account for this characteristic of the labor market, also because we focus on tax changes at high income levels in the present study. If the respondent only participates in the Labor Force Survey in one quarter or if information on actual hours is missing (for example due to illness), contractual hours are used instead. Contractual hours are also used if there is a big difference between contractual and actual working hours.

In order to estimate the conditional logit model, it is necessary to simulate the counterfactual disposable income levels for each discrete alternative, for each individual. Since the Labor Force Survey does not contain any wage information, we compute the hourly wage as yearly wage income (obtained from register-based tax return data) divided by hours per year (measured as 48 times the average weekly hours). The log of computed wage rates is then regressed on individual characteristics using, for females, a Heckman two-stage regression (Heckman, 1979) to account for the selection of individuals not participating (coded as home-working in at least one of the four quarters), while a standard OLS regression is used for males.[28] The number of children and wealth are used as exclusion restrictions under the hypothesis that these variables affect participation, but not wages. Individuals with improbably low or high computed hourly wage rates (under NOK 60 or above NOK 1,200 in 2004) were excluded in the wage regression. For all individuals, we used the predicted wage rate, accounting for a random effect by adding an error term based on draws (30 draws per individual) from a normal

[28] The number of home-working males is small.

distribution.

The actual and counterfactual consumption levels are simulated by multiplying the wage rate by the median point of the discrete intervals. For couples, the income level of the spouse is assumed to be exogenously given and included in non-labor income. As seen in Section 2.2.1, consumption is modeled as $C = f(hw, I)$, where a tax simulation program is used to simulate taxes and disposable income for each individual's hypothetical working hours choice. Tables 2.B1 and 2.B2 report the results of the wage equation regressions, whereas Tables 2.B3 and 2.B4 show the results of the labor supply multinomial logit model. For all four groups we observe positive marginal utility of both consumption and leisure (α_0 and $\beta_0 + \gamma X$ are positive), and α_1 and β_1 are less than 1, which implies that the likelihood functions are strictly concave.

In order to further evaluate the estimation results, Figure 2.B1 shows diagrams of the actual frequencies of working hours and the corresponding probability distribution based on model simulations, for single females, single males, and females and males in couples. The simulated probabilities are derived by calculating the average probability for each choice of hours, based on the individual probabilities. We see that there is close correspondence between observed and predicted choices.

Table 2.B1: Wage regressions for single males and males in couples: log of hourly wage as the dependent variable

	Single males		Males in couples	
	Coefficient	Std error	Coefficient	Std error
Experience	0.0207^{***}	(0.0026)	0.0282^{***}	(0.0023)
Experience squared	-0.0003^{***}	(0.0001)	-0.0005^{***}	(0.0000)
Low education	-0.1150^{***}	(0.0244)	-0.0919^{***}	(0.0169)
High education	0.2379^{***}	(0.0162)	0.2611^{***}	(0.0113)
Residence in densely populated area	0.0864^{***}	(0.0143)	0.1054^{***}	(0.0110)
Non-western origin	-0.1270^{**}	(0.0403)	-0.1395^{***}	(0.0244)
Business code (ref public sector)				
Industry	0.1339^{***}	(0.0182)	0.1561^{***}	(0.0128)
Commerce	0.0130	(0.0192)	0.0827^{***}	(0.0141)
Financial	0.1012^{***}	(0.0218)	0.1556^{***}	(0.0151)
Constant	4.8983^{***}	(0.0305)	4.8425^{***}	(0.0297)
R-square	0.175		0.191	
Number of observations	2,336		4,775	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.B2: Wage regressions (Heckman two-stage selection model) for single females and females in couples: log of hourly wage as the dependent variable

	Single females		Females in couples	
	Coefficient	Std error	Coefficient	Std error
Experience	0.0173***	(0.0021)	0.0136***	(0.0023)
Experience squared	-0.0003***	(0.0000)	-0.0002***	(0.0000)
Low education	-0.0732***	(0.0217)	-0.0778***	(0.0162)
High education	0.2022***	(0.0133)	0.2071***	(0.0105)
Residence in densely populated area	0.0537***	(0.0121)	0.0650***	(0.0098)
Non-western origin	-0.0409	(0.0390)	-0.0292	(0.0246)
Business code (ref public sector)				
Industry	0.1630***	(0.0207)	0.1127***	(0.0144)
Commerce	-0.0088	(0.0149)	0.0094	(0.0114)
Financial	0.0872***	(0.0176)	0.1196***	(0.0136)
Constant	4.8652***	(0.0257)	4.8982***	(0.0328)
Participation				
Experience	0.0886*	(0.0344)	0.0820***	(0.0170)
Experience squared	-0.0018*	(0.0007)	-0.0018***	(0.0003)
Low education	-0.2047	(0.2290)	-0.2783*	(0.1160)
High education	0.9193***	(0.2679)	0.4748***	(0.0926)
Residence in densely populated area	0.0347	(0.1855)	0.0062	(0.0845)
Non-western origin	-0.9852***	(0.2810)	-0.7655***	(0.1323)
Business code (ref public sector)				
Industry	-0.2066	(0.2490)	0.1786	(0.1322)
Commerce	0.3213	(0.2303)	0.2037*	(0.0991)
Financial	-0.1286	(0.2290)	-0.3698***	(0.0952)
Married			0.0007	(0.0993)
No. children under 3 years	-0.9486***	(0.2476)	-0.1609	(0.0984)
No. children under 6 years	-0.1894	(0.2662)	-0.1811	(0.0958)
No. children under 12 years	-0.2978	(0.1556)	-0.2012***	(0.0594)
Net wealth in NOK 10,000	-0.0033*	(0.0014)	-0.0015*	(0.0008)
Constant	1.4808***	(0.3866)	1.3079***	(0.2374)
Mills lambda	-0.1385*	(0.0620)	-0.2312***	(0.0678)
Observations	2,169		4,771	
Censored observations	45		210	
Wald chi2	454.83		755.75	
Prob<chi2	0.0000		0.0000	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.B3: Parameter estimates of the labor supply model, single females and single males

		Single females		Single males	
		Coefficient	Std error	Coefficient	Std error
Consumption					
Constant (Scale 10^{-4})	α_0	0.5754***	(0.0736)	0.7975***	(0.1138)
Exponent	α_1	0.8082***	(0.0624)	0.7368***	(0.0540)
Leisure					
Age	γ_1	-0.0243	(0.0311)	-0.0527	(0.0428)
Age squared	γ_2	0.0005	(0.0004)	0.0008	(0.0005)
High education	γ_3	-0.1377	(0.0773)	0.0801	(0.1269)
Low education	γ_4	0.1213	(0.1430)	-0.1073	(0.1818)
No. children under 6 years	γ_5	-0.5097**	(0.1642)	-0.369	(0.4103)
No. children above 6 years	γ_6	-0.1520*	(0.0738)	-0.5175*	(0.2120)
Residence in densely pop area	γ_7	-0.2573**	(0.0831)	0.1216	(0.1157)
Non-western origin	γ_8	0.223	(0.2222)	0.4157	(0.3343)
Constant (Scale 1/80)	β_0	2.1765**	(0.7622)	3.3612**	(1.2892)
Exponent	β_1	-2.4314***	(0.2602)	-1.5869***	(0.3166)
Opportunity measure: $log\,\theta = f_1 + f_2 S$					
Constant	f_1	-1.1574	(1.0491)		
Years of education	f_2	0.1388	(0.0901)		
Opportunity density of offered hours					
Full-time peak		1.1117***	(0.0547)	1.3801***	(0.0505)
Number of observations		2,208		2,378	

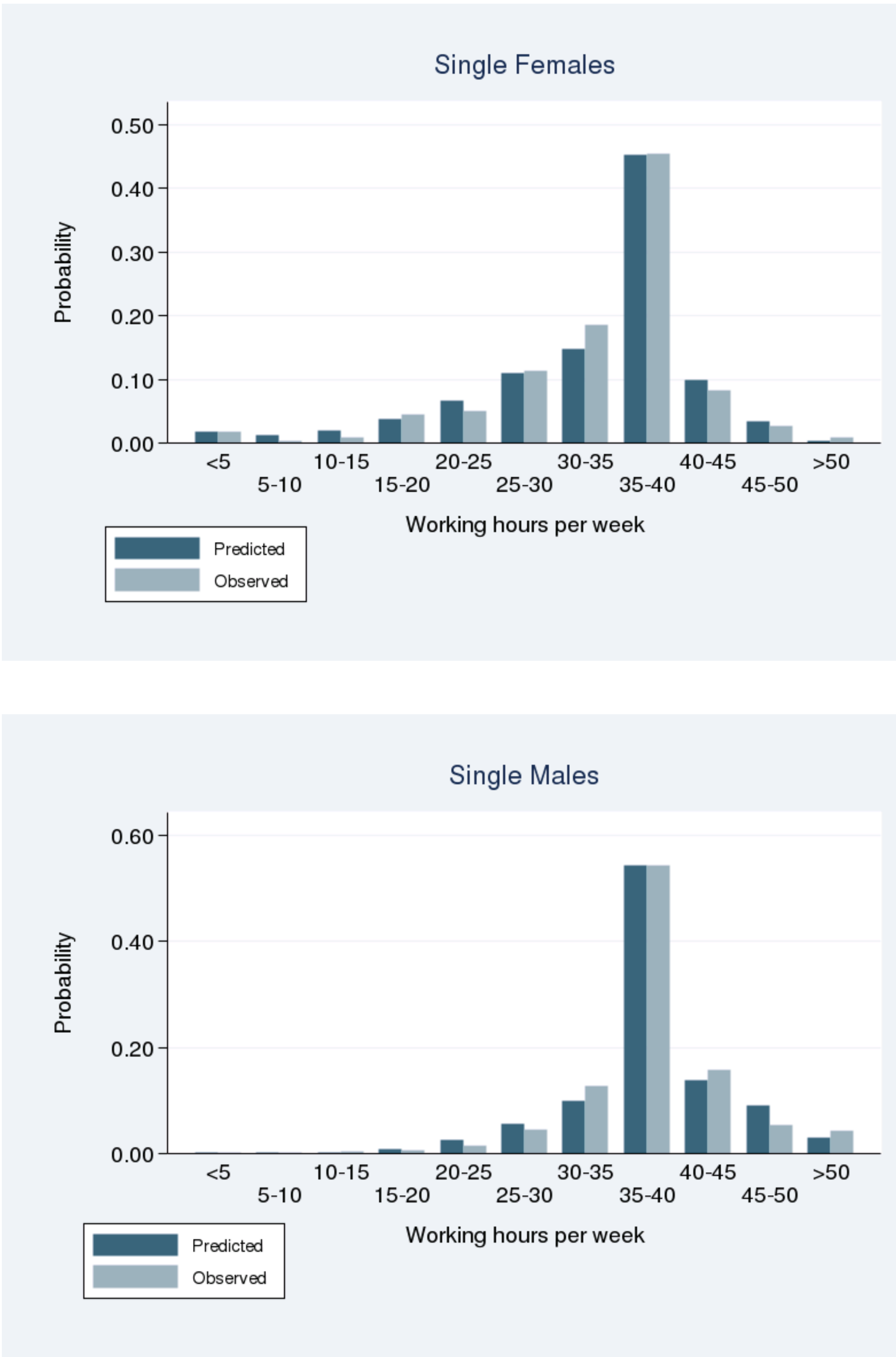
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

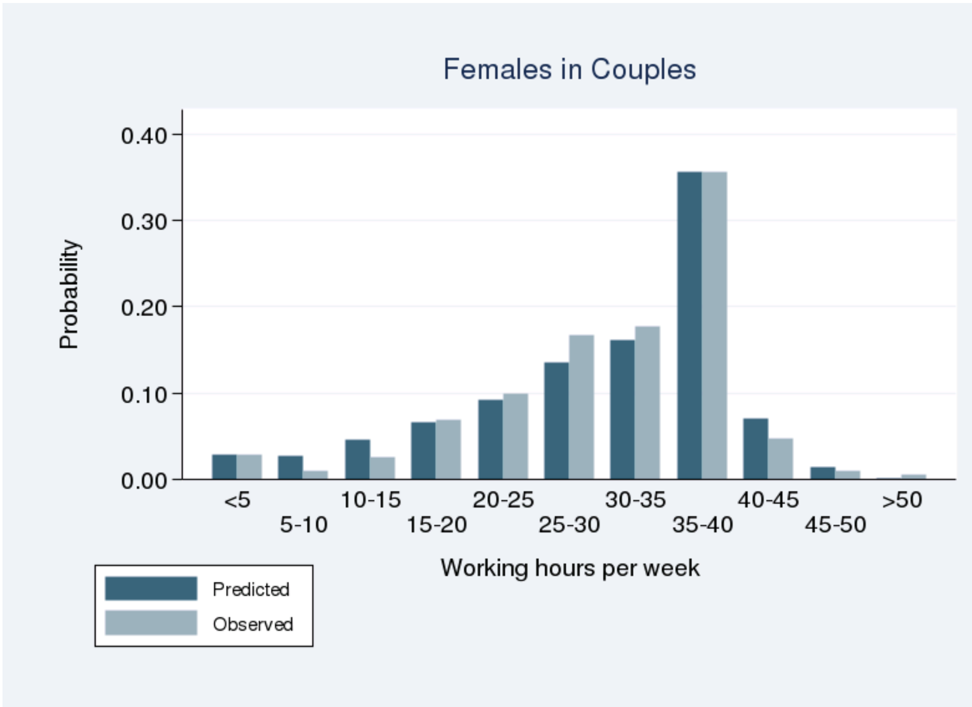
Table 2.B4: Parameter estimates of the labor supply model, females and males in couples

		Females in couples		Males in couples	
		Coefficient	Std error	Coefficient	Std error
Consumption					
Constant (Scale 10^{-4})	α_0	0.7534^{***}	(0.0663)	1.4557^{***}	(0.1543)
Exponent	α_1	0.8844^{***}	(0.0281)	0.7682^{***}	(0.0386)
Leisure					
Age	γ_1	-0.0086	(0.0235)	-0.0402	(0.0363)
Age squared	γ_2	0.0004	(0.0003)	0.0007	(0.0004)
High education	γ_3	-0.2178^{***}	(0.0466)	0.2526^{**}	(0.0922)
Low education	γ_4	0.016	(0.0876)	-0.2817^{*}	(0.1154)
No. children under 6 years	γ_5	-0.0314	(0.0443)	-0.1953^{**}	(0.0706)
No. children above 6 years	γ_6	-0.0119	(0.0296)	-0.2574^{***}	(0.0646)
Residence in densely pop area	γ_7	-0.2206^{***}	(0.0517)	-0.0295	(0.0730)
Non-western origin	γ_8	0.1354	(0.1117)	0.2702	(0.1739)
Constant (Scale 1/80)	β_0	1.5258^{**}	(0.5155)	3.3593^{***}	(0.8972)
Exponent	β_1	-2.7613^{***}	(0.1503)	-1.7919^{***}	(0.1512)
Opportunity measure: $log\,\theta = f_1 + f_2 S$					
Constant	f_1	-2.8891^{***}	(0.5148)		
Years of education	f_2	0.1946^{***}	(0.0449)		
Opportunity density of offered hours					
Full-time peak		0.9405^{***}	(0.0381)	1.4418^{***}	(0.0355)
Number of observations		4,841		4,814	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 2.B1: Predicted and observed probabilities for working hours




Females in Couples
Probability
0.40
0.30
0.20
0.10
0.00
<5
5-10
10-15
15-20
20-25
25-30
30-35
35-40
40-45
45-50
>50
Working hours per week
Predicted
Observed


Males in Couples
Probability
0.60
0.40
0.20
0.00
<5
5-10
10-15
15-20
20-25
25-30
30-35
35-40
40-45
45-50
>50
Working hours per week
Predicted
Observed

Chapter 3

The Dynamics of Earnings Responses to Tax Changes

3.1 Introduction

In the theoretical and empirical literature on taxation and labor supply responses, it is typically assumed that individuals react instantly to changes in the tax schedule, either at the time of implementation or already at the time when the policy reform is announced. Recently, it has been suggested that adjustment or search costs and labor market frictions should not be neglected when analyzing microdata with relatively small changes in the tax schedule (Chetty, 2012; Chetty et al., 2011; Blundell et al., 2011). However, there has been little focus on how optimization frictions affect the dynamics of adjustment in labor supply over time. One exception is Holmlund and Söderström (2011), who differentiate between short and long run responses in labor earnings by a dynamic panel data approach. They argue that neglecting the dynamic nature of individual earnings responses to taxation might not only lead to imprecise predictions regarding timing, but possibly also to a misspecified empirical framework.

In this chapter, I adopt the dynamic panel data framework by Holmlund and Söderström (2011) and compare it with estimates from the conventional static panel approach by Gruber and Saez (2002). I examine earnings responses to changes in the marginal tax rates for median and high income wage earners. High-quality panel data on income and a wide range of individual characteristics for the complete Norwegian population is utilized to estimate the elasticity of earnings with respect to net-of-tax ($1-$marginal tax rate), while controlling for non-tax related aspects. I analyze a recent 14 years period where the surtax schedule (rates and brackets) were substantially altered. This provides exogenous variation in the tax rate through both tax cuts and tax surges over the period.

There are a range of possible explanations for slower earnings responses and consequently the need for a dynamic model framework. One argument is wage earners' lack of information regarding changes in the tax schedule, as people might learn the tax code slowly and not until after the change in disposable income. Holmlund and Söderström (2011) denote this cause a lag in the diffusion of information. They also mention so-called habit persistence and costs associated with changes in hours worked, in addition to responses through human capital investment (or on-the-job training) which leads to a rise in earnings only after some time has passed. The same argument applies to responses in effort and responsibility, which result in higher expected future earnings. Both the lagged response to taxation and the interdependence of previous earnings induce a slower observed pace of adjustment. The autoregressive effect of earnings can be related to so-called state dependence (see Heckman, 1981), which broadly defined reflects changes in preferences, constraints or prices due to past experience (see Chapter 4 for an intertemporal structural model incorporating state dependence). In the present study, I find a small effect from lagged net-of-tax rates, but a strong autoregressive effect in earnings, which causes about twice as large predicted responses in earnings in the long run as opposed to the short run.

To get a better impression of the magnitude of the estimated elasticities in the short and long run, I simulate the revenue implication for the government of a hypothetical change in the surtax rate. The simulation results uncover that although the estimated elasticities seem small, a non-negligible amount of the revenue loss due to cuts in the surtax rate is predicted to be financed through an induced increase in generated income.

This chapter is structured as follows. In Section 3.2, I review the static and dynamic methodological background and literature. In Section 3.3, I describe the tax changes over the period 1995-2008 in light of the institutional settings with the Norwegian tax system. Next, in Section 3.4 the panel data source and the restricted data set is described. In Section 3.5, I present the results for the static three-year panels (Section 3.5.1) and the dynamic panel data model (Section 3.5.2). Next, in Section 3.6, I simulate the effects of a hypothetical reduction in the surtax rate given the estimated elasticities. Finally, in Section 3.7, I provide the conclusion and a short discussion of the results.

3.2 Methodological Background and Literature Review

The literature on elasticity of taxable income, introduced by Lindsey (1987) and Feldstein (1995, 1999) was motivated by capturing the full set of responses and efficiency loss to income taxation in one sufficient statistic. Surprisingly large responses were estimated (elasticities above 1), which suggest considerable efficiency loss due to income taxation. More recent studies find elasticities of taxable income hovering around 0.4, still considerably larger than estimated labor supply elasticities (with some exceptions) for prime age men and high income earners (see surveys such as Meghir and Phillips, 2008 and Saez, Slemrod, and Giertz, 2012). These findings are consistent with the argument in Slemrod (1992, 1995) considering a hierarchy of behavioral responses, in which timing of transactions, avoidance behavior and real decisions of individuals and firms (such as responses in labor supply) can be ordered from the most to the least responsive.

Although it might be desirable to capture the total behavioral responses to income taxation by analyzing taxable income, Slemrod and Kopczuk (2002) argue that the elasticity of taxable income is not a "deep" structural parameter, but depends on institutions and the broadness of the tax base. Moreover, Chetty (2009a) demonstrates that excess burden does not only depend on taxable income elasticity but also on the total earned income elasticity, and recently, Piketty, Saez, and Stantcheva (2011) argue that the responses in taxable income can be divided into the real labor supply responses, the tax avoidance responses and the bargaining responses, where only the first mentioned might be the legitimate measure in an optimal tax formula. As a consequence, it is of prior interest to distinguish between tax avoidance (included income shifting and tax planning) and real responses in earnings (which is the focus of this study).

In the following, I present the theoretical framework behind the literature on elasticity of taxable income with a focus on labor earnings responses. I review the conventional empirical methodology (see e.g. Gruber and Saez, 2002),[1] in order to subsequently outline the alternative dynamic panel specification suggested by Holmlund and Söderström (2011).[2]

The theoretical framework behind the literature on elasticity of taxable income departs from a slightly modified textbook model of labor supply, in which hours of work are replaced by taxable income. Individuals are assumed to maximize a utility function which increases in consumption (C) and decreases in taxable income (q), given a budget constraint where taxable income is taxed with the marginal tax rate τ. The intuition is that not only hours worked are

[1] The conventional approach for estimating the elasticity of net-of-tax rate was introduced by Auten and Carroll (1999) and extended by Gruber and Saez (2002). Several studies have followed this approach also in Europe, for example Aarbu and Thoresen (2001) for Norway, Bækgaard (2010) and Kleven and Schultz (2011) for Denmark, Hansson (2007) for Sweden, Gottfried and Witczak (2009) for Germany and Kiss and Mosberger (2011) for Hungary.

[2] See Massarrat-Mashhadi and Werdt (2012) for a study utilizing a similar dynamic framework to analyze German tax data.

affecting individuals utility negatively, but more generally the effort to acquire income.

$$\begin{aligned} &Max_{c,q}\, U(C,q)\ , \qquad U_C > 0, U_q < 0 \\ &s.t\, C = (1-\tau)q + V\ , \quad V = I + (\tau q - T(q)) \\ &\qquad\qquad\qquad\qquad\qquad\qquad \Rightarrow q = q(1-\tau, V) \end{aligned}$$

Accordingly, this framework can be used for analyzing labor earnings in a Scandinavian institutional setting with a dual income tax system. Following Blomquist and Selin (2010), τ is now the marginal tax rate on labor income, V includes non-labor income net of tax (I) and so-called virtual income $(\tau q - T(q))$ where $T(q)$ are total taxes on labor income q.

The conventional empirical log-log specification in levels can be described as follows.

$$log(q_{it}) = \mu_i + \alpha_t + e \cdot log(1-\tau_{it}) + \gamma \cdot X_{it} + \delta \cdot X_{it} \cdot t + \varepsilon_{it} \tag{3.1}$$

where labor income (q_{it}) in period t depends on an individual specific term (μ_i), a time specific term (α_t), net-of-tax $(1-\tau_{it})$, individual characteristics (X_{it}) and the interaction between individual characteristics and a time or age trend ($X_{it} \cdot t$). The elasticity of earnings with respect to changes in the net-of-tax rate is defined by $e = \frac{1-\tau}{q}\frac{dq}{d(1-\tau)}$. In most studies the income effect through virtual income (V_t) is ignored, with the presumption that income effects are small.[3]

After differencing (when all individual characteristics are treated as time-invariant) equation (3.2) is obtained.

$$log(\frac{q_{i,t+s}}{q_{i,t}}) = \alpha_{t,t+s} + e \cdot log(\frac{1-\tau_{i,t+s}}{1-\tau_{i,t}}) + \beta X_i + (\varepsilon_{it+s} - \varepsilon_{it}) \tag{3.2}$$

As described in Chapter 2, the net-of-tax rate change is clearly endogenous and is instrumented in the conventional literature by a tax rate change for a constant (inflation-adjusted) initial income level $q_{it}(1+g)$.

$$log(\frac{1-\tau_{t+s}(q_{it+s})}{1-\tau_t(q_{it})}) \Leftarrow log(\frac{1-\widehat{\tau_{t+s}}(q_{it}(1+g))}{1-\tau_t(q_{it})})$$

The problem of distinguishing tax rate responses from mean reversion and trends in the income distribution (see Chapter 2) is accounted for by including a function of income in period

[3] Gruber and Saez (2002) proposed a method for including income effects and find that the income effects are minor using US data. I include income effects in some specifications presented in Section 3.5, following the approach by Blomquist and Selin (2010).

t, $f(q_{it})$, in order to eliminate the correlation between the error term $(\varepsilon_{it+3}-\varepsilon_{it})$ and the constructed instrument for the tax rate change. In the baseline regressions in Section 3.5.1, I define $f(log(q_t))$ as a third degree polynomial.[4]

$$log(\frac{q_{it+s}}{q_{it}}) = \alpha + e \cdot log(\frac{1-\tau_{it+s}}{1-\tau_{it}}) + \beta X_{it} + \phi f(log(q_{it})) + (\varepsilon_{it+s}-\varepsilon_{it}) \tag{3.3}$$

Following the conventional approach, as in Chapter 2, I set $s = 3$, stack all available three-year differences in one regression and apply two-step-least-square (2SLS) to estimate equation (3.3). The results are provided in Section 3.5.1.

The choice of three-year spans is motivated in the literature by the recognition that individuals might need some time to react to tax changes. However, there are not many attempts to investigate *when* one can expect individuals to respond and how the earnings responses develop over time.

To focus on the time dimension of tax responses, I follow the alternative approach by Holmlund and Söderström (2011) and estimate a dynamic panel data model.

Holmlund and Söderström (2011) suggest a dynamic panel data model represented by an autoregressive distributed lag model. This represents an extension of the conventional framework in two respects. First, it allows for earnings response to lagged tax rate changes and second it allows for an autoregressive effect in earnings, in which last period affect current earnings.

[4] Some recent studies are critical to the tax instruments, which are not clearly exogenous, and the simple correction through the function of base year income. The criticism is mainly based on that one cannot be sure that the error term is uncorrelated with the tax instrument. One way of seeing the conventional approach is that one includes a function of base year income to extract the part of the error term $(\varepsilon_{it+s}-\varepsilon_{it})$ which is correlated with the tax instrument (since the tax instrument is a function of the base year income and therefore correlated with ε_{it}). One must further assume that income in period t is uncorrelated with the error term in period $t+s$. This statement is criticized by Blomquist and Selin (2010) and Weber (2011), who point out that there might be serial correlation in the error term. Weber (2011) investigates this by constructing instruments based on one period prior to the initial period $(t-1)$ in addition to the conventional instruments constructed by the base year t. By testing the over-identified equation, she finds that the conventional instrument is not exogenous, which leads to biased results. She obtains three times as large estimates for the elasticity of taxable income by using the alternative approach utilizing the same data set as in Gruber and Saez (2002). The drawback is however that the estimates have larger standard errors and the F-statistics falls dramatically, pointing to that the instruments are considerably weaker. Blomquist and Selin (2010) base their instrument on a period in the middle (between period t and $t+s$) as they assume the correlation with ε_{it} and $\varepsilon_{i,t+s}$ to be similar and cancel out. They base their approach on a long period panel and attain relatively large (however, imprecisely measured) elasticities both of hourly wage rates and taxable income.

After first differencing the dynamic panel data specification in levels, one obtains

$$log(\frac{q_{it+1}}{q_{it}}) = \kappa log(\frac{q_{it}}{q_{it-1}}) + e_1 log(\frac{1-\tau_{it+1}}{1-\tau_{it}}) + e_2 log(\frac{1-\tau_{it}}{1-\tau_{it-1}}) + \beta X_{it} + (\varepsilon_{it+1} - \varepsilon_{it}) \quad (3.4)$$

where the conventional static specification in equation (3.2) is a special case of the dynamic framework with $\kappa = e_2 = 0$. In the dynamic model, e_1 measures the short run elasticity whereas the long run elasticity is given by the expression $(\frac{e_1+e_2}{1-\kappa})$ (See e.g. Bond, 2002).

As the autoregressive term in the dynamic equation is endogenous by construction (q_{it} and ε_{it} are correlated), it is instrumented according to standard methods in the dynamic panel literature. Holmlund and Söderström (2011) applies the IV approach of lagged income level proposed by Anderson and Hsiao (1981).

$$log(\frac{q_{it}}{q_{it-1}}) \Leftarrow log(q_{it-1})$$

Both the contemporary and lagged change in net-of-tax rate need to be instrumented. The instrument is now constructed based on earnings in period $t-1$ (instead of period t in the conventional approach), where $g_{t-1,t+1}$ refers to growth in median income from period $t-1$ to period $t+1$.

$$log(\frac{1-\tau_{it+1}}{1-\tau_{it}}) \Leftarrow log(\frac{1-\tau_{t+1}((1+g_{t-1,t+1})q_{it-1})}{1-\tau_t((1+g_{t-1,t})q_{it-1})})$$

$$log(\frac{1-\tau_{it}}{1-\tau_{it-1}}) \Leftarrow log(\frac{1-\tau_t((1+g_{t-1,t})q_{it-1})}{1-\tau_{t-1}(q_{it-1})})$$

In Section 3.5.2, I present the results of the dynamic IV specification in addition to an alternative GMM-difference specification.

3.3 Institutional Background: Income Taxation in Norway

Norway has a dual income tax system characterized by a flat tax on capital income combined with a step-wise progressive tax on labor income. This is implemented through two separate tax bases denoted “ordinary income” and “personal income”. Ordinary income consists of all taxable income from labor and capital after basic allowance and other tax deductions are subtracted, and is taxed with a flat rate of 28 percent. Personal income includes gross labor income and serves as the tax base for social security contribution[5] and a two tier surtax rate system.

Two separate tax classes apply to Norwegian residents. The majority are taxed individually in tax class 1 (about 90 percent). Single parents are taxed in tax class 2, in which both the basic allowance and the bracket point for the first step surtax rate (in most years) are extended. Tax class 2 also applies to married couples if joint assessment is beneficiary (usually only one-income families). There are no additional local taxes which differ across municipalities. However, for individuals with residence in Northern Troms and Finnmark (North in Norway) an additional tax deduction and lower rate for the first surtax tier applies.

Considerable changes in the surtax schedule for labor income provide exogenous variation in tax rates over the period. In 1995-1998 the two surtax rates were respectively 9.5 percent and 13.5 percent. Both brackets were implemented at a relatively low income level, according to the respectively 60 and 70th percentile of the wage earner sample (sample restrictions are described in Section 3.4). In 1999, the two tier brackets were temporarily combined to one with a tax rate equal to 13.5 percent followed by a new surtax system in the period 2000-2004 with tax rates of 13.5 percent and 19.5 percent. The second step of the surtax system was now, however, only levied on particular high levels of labor income corresponding to the top 1-2 percent of the wage earner sample. Related to the more comprehensive 2006 tax reform (as described in Chapter 2), the surtax rates were considerable reduced to 9 and 12 percent in the most recent period 2006-2008.[6] In the new surtax schedule, the brackets were again scaled down, such that individuals with income in percentile 92 or above were assigned to the second surtax tier.

Figure 3.1 depicts the statutory marginal tax rate per income percentile for wage earners in the regular first tax class over the period 1995-2008. Note again that the percentile measure is based on the sample of wage earners described in Section 3.4. To make the figure more tractable I have divided into three main periods as described above and used the average threshold percentile in each period. The transition years 1999 and 2005 are excluded from the figure (although all the years are included in the empirical analysis).

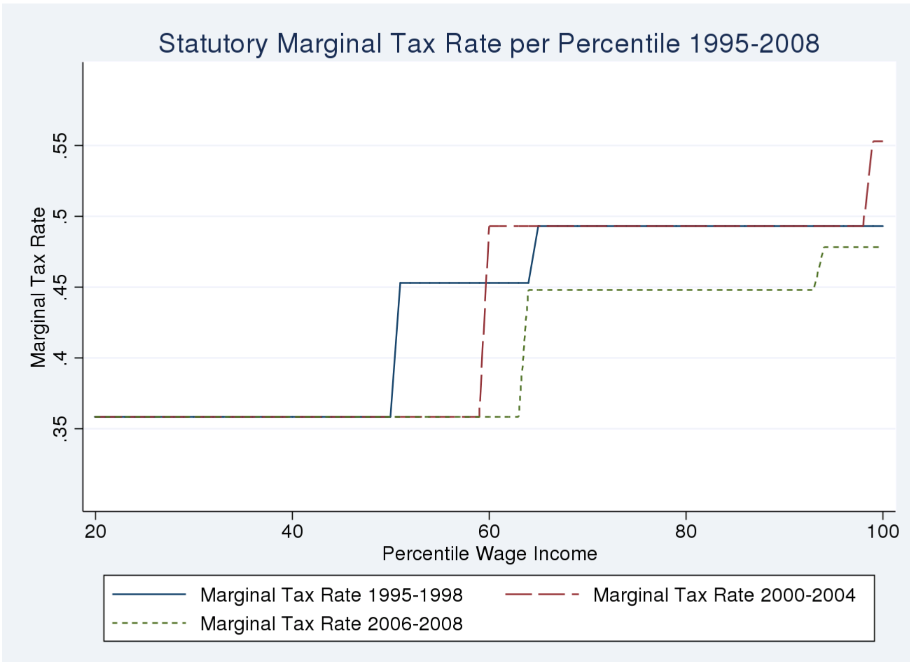
Note that it is advantageous for identification of the tax responses that the highest income earners experienced both positive and negative changes in the tax rates over the period. In

[5] Social security contributions add up to 3.0 percent for pensions, 7.8 percent for wage earners and 10.2/11 percent for self-employed

[6] 2005 was a transition period with rates of 12 and 15 percent.

addition, individuals in the second tax class or with residence in the north of Norway provide additional exogenous variation across income levels.

Figure 3.1: Statutory marginal tax rate for Norwegian wage earners 1995-2008



3.4 Data

The data source is a register data set collected from tax return and income statistics which covers the complete Norwegian population. The data set includes detailed information on income, wealth and individual characteristics such as gender, age, education level and field of education, marital status, residence and origin. A family or household number makes it possible to identify spouse and children. In addition, each individual is coded with a personal identification number, such that the panel dimension easily can be exploited. Detailed information on the data source is described in Statistics Norway (2005).

I restrict the data set to wage earners in the age group 25-62 over the period 1995-2008. Wage earners are defined to have labor income as their main source of income. Students and individuals receiving pensions or unemployment benefits are excluded. In addition I exclude individuals with any self-employed income and individuals with very high capital income (the highest percentile) or with negative capital income.[7] The reason for these exclusions is to avoid various bias to the results, as filers with extreme positive or negative wage growth for other reasons than taxes might impact the results. It is, however, important to be aware of that when only considering strictly defined prime age wage earners; the responses to taxation might be lower than for special groups with a more fragile relation to the labor market. The tax responses for these special groups need, in my view, to be studied separately and is not within the scope for this study. Note, however, that in contrast to e.g. Holmlund and Söderström (2011), I do not constrain the panel to be balanced over the complete period.

The marginal tax rates and total taxation are simulated based on information on taxable income, municipality, tax class and deductions.[8] Instead of using the strictly defined marginal tax rate based on a unit increase in income, I simulate (the more relevant) marginal tax rate by increasing labor income by five percent.

General summary statistics are provided in the Appendix.

3.5 Results

In the following, I first present the results for the static conventional panel as a reference (Section 3.5.1) and compare with results from the dynamic panel data approach (Section 3.5.2).

3.5.1 The Static Panel Data Model

I apply the conventional differenced static panel specification described in equation (3.3) in Section 3.2, and stack the ten three-year differences of 1995-1998, 1996-1999, ... , 2004-2007, 2005-2008.

The tax rate change is instrumented by a tax rate change for an inflation-adjusted initial income level $q_t(1+g)$, initial tax class and residence, where g corresponds to the growth in median labor income over the period t to $t+s$.[9]

[7] The age and income restrictions apply for each year, such that when considering three-year panels, a wage earner is included in the regression sample if both years fulfill the restrictions.

[8] Note that deductions are not relevant for the surtax rate, but for the flat tax on total income (follows from the dual income tax system).

[9] The median income earner was not affected by the tax rate changes. Moreover, the results are robust to an alternative method to predict growth rates based on the individual's education and age group. All individual characteristics are also included in the first step regression for the tax rate change, so it should not be crucial how the income level is inflated.

Individual characteristics, X_{it}, include dummy for marital status, gender, residence in Oslo and non-western origin, years of education, age, age squared, wealth, field of education and year dummies. Number of newborn children (under 3 years), number of children under 6 and number of children between 6 and 18 are based on children's age in year $t+s$.

As in Chapter 2, I restrict the regression analysis to individuals in earnings percentile 33 or above in the base year t (no restriction on labor income in year t+s). The reason is that the mean reversion problem is especially severe for individuals with initially low income levels, which makes this group less appropriate as a control group. Moreover, the focus is on responses to surtax rates, which only affected about the upper one third of the income distribution for wage earners.

In order to check the influence of allowing for income effects, I include virtual income in one specification following the approach by Blomquist and Selin (2010). The rationale for including virtual income is that it represents an income effect which according to theory would have a negative impact on working hours and effort. Remember that virtual income (V_{it}) can be described as

$$V_{it} = I_{it} + (\tau_{it} q_{it} - T_{it}(q_{it}))$$

where non-labor income (I_{it}) is defined as capital income and transfers net of tax in addition to spouse's income if in couple.

The full model with virtual income is described as

$$log(\frac{q_{t+s}}{q_t}) = \alpha + e \cdot log(\frac{1-\tau_{t+s}}{1-\tau_t}) + \vartheta \cdot log(\frac{V_{t+s}}{V_t}) + \beta X_{it} + \phi f(log(q_t)) + (\varepsilon_{it+s} - \varepsilon_{it}) \quad (3.5)$$

There are two sources of endogeneity present for $log(\frac{V_{i,t+s}}{V_{i,t}})$. First, it might be a problem of reversed causality if labor income affects capital and transfers. $I_{i,t+s}$ is therefore instrumented by using non-labor income in period t and inflating this to period $t+s$ by using median growth in total income. Second, as total taxes are directly dependent on labor income growth, the same endogeneity problem occurs as with the net-of-tax rate. I therefore again instrument marginal tax rate and total taxes by inflating base year labor income by median growth and applying the t+s tax schedule.

$$\widehat{V_{t+s}} = I_t(1+g_{t,t+s}) + (\tau_{t+s}(q_t(1+g_{t,t+s})) \cdot (q_t(1+g_{t,t+s})) - T_{t+3}(q_t(1+g_{t,t+s})))$$

$$log(\frac{V_{t+s}}{V_t}) \Leftarrow log(\frac{\widehat{V_{t+s}}}{V_t}), log(V_t)$$

The main static panel results are presented in Table 3.1. Regression 1 presents the results of the conventional preferred regression whereas regression 2 includes additionally a control for virtual income. The (uncompensated) net-of-tax elasticity is measured to about 0.06 and is slightly reduced when allowing for income effects, which is small and negative as expected, and not significantly different from zero. The individual specific control variables seem to have appropriate signs and sizes. For example is labor income growth higher for male and increasing with education. There is as expected a negative effect on income growth of getting a child, but the income level seems to increase again through higher growth rates as the child(ren) grow older.

A problem with the income elasticity instrument is that it is highly correlated with the net-of-tax instrument, so it seems difficult to distinguish between the two effects especially when looking at sub-groups of individuals. Moreover, to follow the standard approach without income effects in the literature, I choose regression 1 without the income elasticity as a baseline for the robustness tests and for the simulations in Section 3.6.

Although the conventional approach does not consider dynamic responses, it is still indirectly possible to have a look at the timing of responses by regarding different time span of the differenced earnings equation. I report how the measured elasticities depend on the time frame s in Table 3.2. The choice of $s = 3$ in the baseline regressions is standard in the literature, but clearly ad hoc. It seems from these results that the measured elasticities are relatively robust to different time frames. Nevertheless, the shortest time frame provides somewhat lower estimates and it seems like the maximum response is already reached at $s = 3$ or $s = 4$. This is in line with the results in Section 3.5.2 that long run responses are somewhat larger than the immediate effect.

Table 3.1: Three-year static panels: Baseline regressions

	Reg 1		Reg 2	
Net-of-tax elasticity	0.0630***	(0.0021)	0.0559***	(0.0032)
Income elasticity			-0.0015	(0.0013)
Male	0.0467***	(0.0002)	0.0378***	(0.0002)
Age/100	-0.0005	(0.0084)	-0.0400***	(0.0092)
Age squared/10,000	-0.1149***	(0.0101)	-0.0738***	(0.0108)
Married	0.0102***	(0.0002)	0.0081***	(0.0002)
No. newborn children	-0.0589***	(0.0003)	-0.0496***	(0.0003)
No. children under 6	0.0176***	(0.0002)	0.0157***	(0.0002)
No. children above 6	0.0068***	(0.0001)	0.0071***	(0.0001)
Non-western origin	-0.0329***	(0.0005)	-0.0331***	(0.0005)
Residence in large city	0.0112***	(0.0002)	0.0097***	(0.0002)
Years of education	0.0137***	(0.0000)	0.0129***	(0.0000)
Wealth	-0.0004***	(0.0000)	-0.0003***	(0.0000)
Polynomial of base year income	Yes		Yes	
Field of education (dummies)	Yes		Yes	
Year dummies	Yes		Yes	
Constant	0.0019	(0.0021)	0.0276***	(0.0024)
Observations	9,027,613		7,973,036	

Note: Standard errors in parentheses. All wage earners in base year income percentile 33 or above are included in the regressions. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.2: Alternative time frame

Alternative time frame	s=1	s=2	s=3	s=4
Net-of-tax elasticity	0.0440***	0.0551***	0.0630***	0.0622***
	(0.0020)	(0.0020)	(0.0021)	(0.0023)
Observations	11,826,958	10,321,980	9,027,613	7,876,656

Standard errors in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.5.2 The Dynamic Panel Data Model

In the following, I present results for the alternative dynamic specification described in Section 3.2.

For convenience, I follow Holmlund and Söderström (2011) and estimate an equivalent version of (3.4).[10] The full model with income effects is now given by

$$log(\frac{q_{t+1}}{q_t}) = \kappa \, log(\frac{q_t}{q_{t-1}}) + e_1 \left(log(\frac{1-\tau_{t+1}}{1-\tau_t}) - log(\frac{1-\tau_t}{1-\tau_{t-1}}) \right) \tag{3.6}$$

$$+(e_1+e_2)\, log(\frac{1-\tau_t}{1-\tau_{t-1}}) + \vartheta_1 \, log(\frac{V_{t+1}}{V_{t-1}}) + (\vartheta_1+\vartheta_2)\, log(\frac{V_t}{V_{t-1}}) + \beta X_{it} + (\varepsilon_{it+1} - \varepsilon_{it})$$

In addition to the Anderson-Hsiao IV approach with lagged levels described in Section 3.2, I also present results for the Arellano-Bond two-step GMM-difference estimation, which is an extension of the simple IV technique in which all available lagged income levels and first-differences are used as instruments.

$$log(\frac{q_t}{q_{t-1}}) \Leftarrow \begin{array}{lll} log(q_{t-1}) & log(q_{t-2}) & \ldots \\ log(\frac{q_{t-1}}{q_{t-2}}) & log(\frac{q_{t-2}}{q_{t-3}}) & \ldots \end{array}$$

The GMM estimator also takes the resulting structure of the error terms into consideration under the assumption that the level error terms are i.i.d. In theory, the GMM estimator gains efficiency, whereas both the IV method and GMM method should provide consistent results. In practice, the point estimates differ somewhat. For both the IV and GMM estimations, I conduct the Arrelano-Bond test for autocorrelation (AR-2 test statistic).[11]

The analysis is now restricted to individuals with pre-base year income (period $t-1$) in percentile 33 or above. Income shifting is considered by including a variable for the change in capital taxation over the period times the logarithm of capital income in the base year period t: $\triangle\tau_{K,t,t+s} \cdot log(capinc_{it})$. The idea is that an increase in the marginal tax rate for capital induces individuals to shift some of their income to labor income, and the effect is assumed

[10] When estimating (e_1+e_2) it is straightforward to apply the delta-method to obtain standard errors for the long run estimate.

[11] If the original error terms in the level equation are i.i.d., then the differenced model produces error terms with an MA-1 structure. The test for AR-1 process is therefore expected to be positive. More interestingly, it is desirable to test for AR-2 process in the error terms which would predict that the original error terms are autocorrelated. In that case the instruments should be constructed based on further lags.

to increase with the individuals initial capital income. The change in marginal tax rate of capital income is not individual specific (since a flat tax rate), hence it does not suffer from endogeneity problems.

I present results for two models with adequate AR-2 test statistics. First the standard IV approach suggested by Holmlund and Söderström (2011), and next a GMM difference approach with instruments extended to include a third degree polynomial of lagged income levels. The results are reported in Table 3.3. Regression 1 and 2 provide the results for the standard Anderson-Hsiao level instrument, whereas regression 3 and 4 represent the GMM difference specification. Analogous to the baseline estimation for the conventional static model, I report results both with and without income effects. The two models (IV and GMM) provide similar results. The income elasticities are measured to be small negative and slightly reduce the net-of-tax elasticity estimates, similar to the static panel results. The lagged growth in income has a considerably large impact on growth in labor income with an coefficient above 0.5.[12] The standard Anderson-Hsiao IV approach is used as a basis for robustness checks reported in the Appendix and for the simulation exercise conducted in Section 3.6.

The long run elasticities are computed by $(\frac{e_1+e_2}{1-\kappa})$ as described in Section 3.2, and the standard errors are attained by the delta method. The short and long run elasticities and the AR-2 test statistic for the four models above are summarized in Table 3.4.

The dynamic estimations seem to predict that the earnings responses have an immediate effect of 45-65 percent, and increase to close to the full effect after four-five years. The results from the dynamic model predict somewhat larger responses than in the conventional approach, as the static three-year panel approach should be compared to an average of the first, second and third year responses in the dynamic model.[13] Both models, however, suggest that the real responses for high income wage earners are modest.

[12] The large coefficient is partly due to the income restriction (above percentile 33) in year $t-1$. Under this restriction income growth is considerably less volatile

[13] This because in the static panels one does not distinguish between tax changes occurring in the beginning and in the end of each time-span. An average of the first, second and third year responses corresponds to an elasticity of 0.06-0.09.

Table 3.3: Dynamic model: Main results

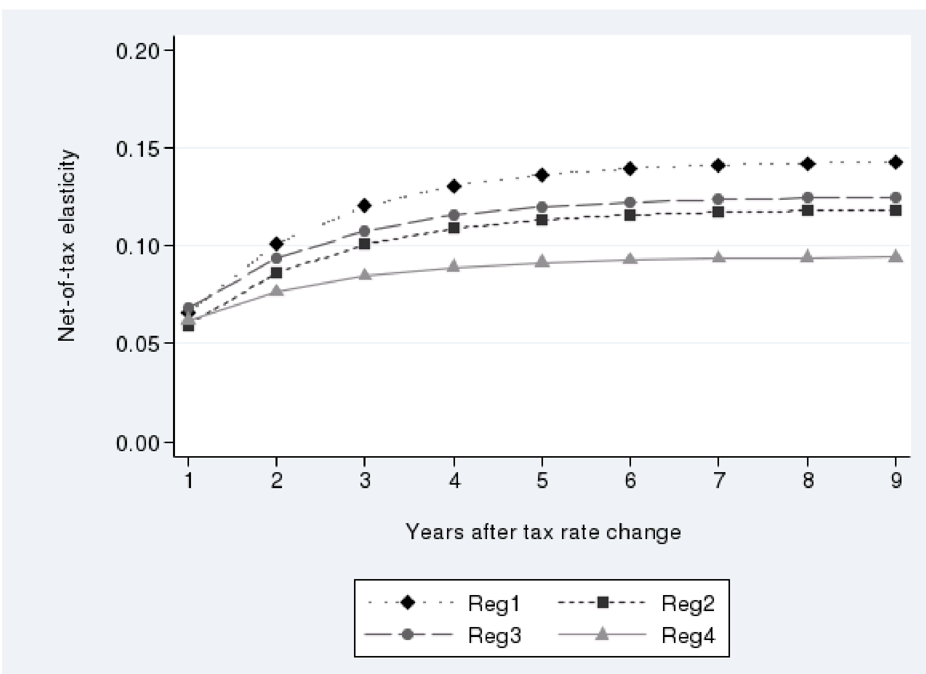
	IV-2SLS		GMM-dif	
	Reg 1	Reg 2	Reg 3	Reg 4
Net-of-tax elasticity (e_1)	0.0661***	0.0598***	0.0681***	0.0620***
	(0.0061)	(0.0062)	(0.0056)	(0.0057)
Sum net-of-tax elasticities ($e_1 + e_2$)	0.0653***	0.0532***	0.0562***	0.0423***
	(0.0054)	(0.0055)	(0.0047)	(0.0048)
Income elasticity (ϑ_1)		-0.0025***		-0.0020***
		(0.0004)		(0.0004)
Sum income elasticity ($\vartheta_1 + \vartheta_2$)		-0.0046***		-0.0055***
		(0.0005)		(0.0005)
Income shifting control		0.0029***		0.0033***
		(0.0002)		(0.0001)
Lagged income growth (α)	0.5444***	0.5532***	0.5513***	0.5529***
	(0.0051)	(0.0051)	(0.0044)	(0.0044)
Male	0.0001	-0.0001	0.0002	0.0002
	(0.0002)	(0.0002)	(0.0001)	(0.0001)
Age/100	0.0307***	0.0225**	0.0098	0.0007
	(0.0072)	(0.0074)	(0.0054)	(0.0056)
Age squared/10,000	-0.0534***	-0.0451***	-0.0258***	-0.0189**
	(0.0084)	(0.0086)	(0.0061)	(0.0063)
Married	0.0022***	0.0018***	0.0020***	0.0018***
	(0.0001)	(0.0002)	(0.0001)	(0.0001)
Newborn*Male	-0.0047***	-0.0044***	-0.0050***	-0.0045***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Newborn*Female	-0.0519***	-0.0511***	-0.0516***	-0.0513***
	(0.0004)	(0.0005)	(0.0006)	(0.0006)
No. children under 6	0.0038***	0.0039***	0.0038***	0.0037***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
No. children above 6	0.0013***	0.0013***	0.0012***	0.0012***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Non-western origin	-0.0078***	-0.0075***	-0.0074***	-0.0068***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Residence in large city	-0.0020***	-0.0020***	-0.0021***	-0.0020***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Years of education	0.0011***	0.0010***	0.0011***	0.0010***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Wealth	0.0002***	0.0001***	0.0002***	0.0002***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Constant	0.0087***	0.0185***		
	(0.0018)	(0.0019)		
Observations	8,572,918	8,533,971	8,572,918	8,531,919

Table 3.4: Short and long run elasticities

	IV-2SLS		GMM-dif	
	Reg 1	Reg 2	Reg 3	Reg 4
Short run elasticity	0.0661***	0.0598***	0.0681***	0.0620***
	(0.0061)	(0.0062)	(0.0056)	(0.0057)
Long run elasticity	0.143***	0.119***	0.1253***	0.0945***
	(0.0104)	(0.0103)	(0.0089)	(0.0089)
AR-2 statistic	-2.72	-0.47	-1.19	-0.25
(p-value)	(0.0064)	(0.6357)	(0.2350)	(0.8030)

Note: Standard errors in parentheses.

Figure 3.2: Predicted elasticities over time



3.6 Simulations of a Hypothetical Surtax Cut

In this section, I provide simulation exercises to illustrate what the estimated elasticities infer about the importance of behavioral responses to tax revenues. I compare predictions based on the static model with predictions from the dynamic model (IV-2SLS).[14]

When analyzing to which degree hypothetical tax cuts in surtax rates are self-financed through increased generated income, both the surtax bracket points and the existing tax level are important factors. I consider a tax schedule with only one surtax bracket for simplicity. In the first simulation exercise, the surtax bracket is implemented already at the level of median labor income. In the second and third simulation exercise, it is implemented at a level of two and three times the median income respectively.[15] For each of the three tax schedules, I consider a hypothetical tax reform where the top tax rate is reduced from 50 percent to 45 percent. I simulate the predicted wage growth for each individual based on the estimated regressions, in which the change in net-of-tax rate is altered according to the applicable surtax schedule. The earnings responses to a change in net-of-tax are assumed to be equally distributed among individuals. In each case I simulate the predicted percentage of tax losses which would be financed through additional generated income.[16] A hundred percent would correspond to a self-financed reform, where the mechanical tax losses are exactly offset by the increase in tax revenues due to increased generated income.

The simulation results are summarized in Table 3.5.[17] The static model results infer that between 18 and 26 percent of the tax revenues are returned through increased generated income. The figures are similar in the short run according to the dynamic model. In the long run the

[14] For simplicity, I ignore income effects.

[15] In the current Norwegian surtax schedule the first tax bracket is implemented at an income level of about 1.4 the median income level, and the second tax bracket at a level corresponding to about 2.3 times the median income level.

[16] An example might be clarifying. Imagine a tax schedule with only one surtax bracket. Pre-reform, the marginal tax rate adds up to 50 percent for income above the level of NOK 800,000. Post-reform, the marginal tax rate is reduced to 45 percent whereas the bracket cut-off is retained. Consider an individual earning NOK 1,000,000 before the tax reform. If his elasticity of labor income with respect to net-of-tax is 0.05, he will now choose to earn NOK 1,005,000 (since $log(1-0.5)-log(1-0.45) \approx 0.1$ predicts labor income growth to be (additionally) $0.05 \cdot 10\% = 0.5\%$). It follows, that for this individual, the government has a mechanical loss of $200,000 \cdot 5\% = 10,000$ due to the tax cut. As a consequence of the individual's behavioral change, the government gain extra revenues for the additional 5000 that are now earned according to $5,000 \cdot 45\% = 2,250$. In other words, 22.5 percent of the tax cut losses are returned through increased generated income in an economy with only this individual.

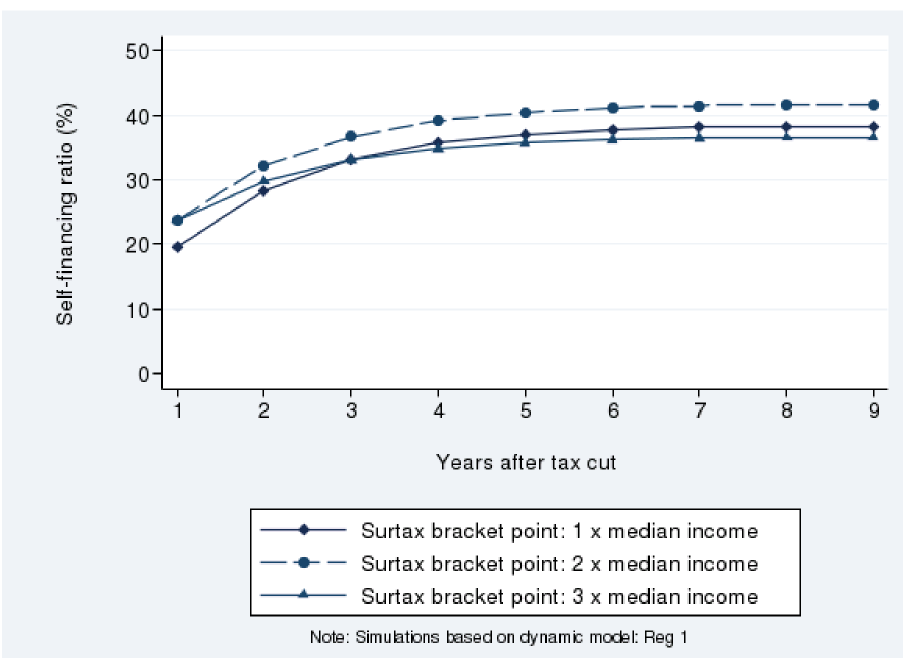
[17] Additional generated income would also have an effect on pay roll taxes and on indirect taxation, so the figures in Table 3.5 can be considered as lower bounds.

dynamic model predicts a self-financing ratio of about 40 percent, which reflects the higher estimated long run elasticities.

Table 3.5: Simulation of hypothetical tax reforms

	Top tax rate reduction from 50% to 45%		
	Self-financing ratio		
	Static model	Dynamic model	
Top tax bracket begins at	Short/long run	Short run	Long run
Median Income	17.6%	19.6%	38.3%
2 x Median Income	23.0%	23.8%	41.7%
3 x Median Income	26.2%	23.7%	36.6%

Figure 3.3: Self-financing ratio over time



3.7 Conclusion and Discussion

There are a number of reasons for why wage earners should not be expected to react instantly to changes in the tax schedule. Partly because of lagged responses to tax rate changes and partly because of an interdependence in earnings over time in which individuals seems to slowly adjust to a new budget constraint over time. There are so far few attempts to analyze the dynamics of earnings responses, although wrongly assuming immediate adjustment can lead not only to imprecise predictions with regard to timing, but also to a misspecified framework.

In this study, I have used high-quality Norwegian register data for the complete population of wage earners to analyze a 14 year period with considerable changes in the surtax schedule (including both tax cuts and tax hikes) for labor income in order to compare a dynamic panel data model proposed by Holmlund and Söderström (2011) with the conventional static panel approach by Gruber and Saez (2002).

The results suggest that long run responses are about twice as high as short run responses and that the conventional method might slightly underestimate results also in the short run. It should however be noticed that, in contrast to other studies on alternative ETI methods, I do not find a strong contradiction between the conventional and the alternative (in this case dynamic) model. It is reassuring that both methods provide similar and relatively small responses in earnings. However, although the estimates are small, I demonstrate that the earnings responses should not be neglected. Additionally conducted simulation results suggest that up to 40 percent of the financing of tax cuts are returned through induced increased income in the long run.

The focus on responses in earnings is interesting as these ideally are more dependent on individual preferences and less on institutions and therefore more comparable over countries, although the responses might to some degree capture non-real bargaining responses as described in Piketty, Saez, and Stantcheva (2011). As income mobility is relatively low for wage earners in the mid and upper part of the income distribution (in contrast to e.g. self-employed), the methodological framework with tax instruments based on a constant income level seem appropriate. The use of a data source with the unique combination of tax returns coupled with socioeconomic characteristics and in particular information on education at the individual level, allows me to control for a wide range of alternative sources for income growth, for instance due to increased globalization and technological change (see e.g. OECD, 2011).

A possible reason for the relatively small measured real responses could be due to that the tax changes are relatively small and vary rapidly over time, such that it is not optimal for a worker to actually prolong education or shift job because of adjustment costs. Notice that although the dynamic panel approach takes into account possible lagged responses and the interdependence in income over time, it does not incorporate expectations about future tax level. To the extent at which tax changes might have been announced or expected before they actually occurred, it might also be a general expectation about whether the current tax level is below or above a sustainable long-run level. This could lead to intertemporal substitution in which one work more in periods with low tax rate. However, as investigated by Blundell et al. (2011), the intertemporal substitution effect might be dominated by labor market frictions inducing people to adjust to the expected future level at the expense of short term optimization.

Appendix

In Table 3.A1 general summary statistics for the wage earner sample is reported. The sample is divided into three periods of 1995-1999, 2000-2004 and 2005-2008. All income are measured in nominal NOK ($1\,Euro \approx 8\,NOK$). Next, in Table 3.A2 summary statistics for the individual characteristics are reported.

Three-year differences in income, net-of-tax, virtual income and the respectively instruments in addition to the income shifting control are reported in Table 3.A3. In Table 3.A4 the three-year differences in net-of-tax instrument and labor income growth is compared for groups of wage earners according to percentiles of base year income.

I expect that individuals with a positive exogenous net-of-tax rate change have incentives for higher growth in labor income. However, due to mean reversion (or drifts in the income distribution) it seems evident that higher base year income is associated with lower labor income growth. According to the figures most of the exogenous (instrumented) tax changes occurred for percentile 100 and least for percentile 33-54. In percentile 55-74 changes are mainly due to altered lower surtax bracket points. Percentile 75-94 experienced no changes in the first periods, but faced lower tax rates in the latter periods, percentile 95-99 accordingly, however with smaller tax rate changes. The most sizable changes occurred during the last periods. Percentile 100 experienced both the largest hikes and cuts in marginal tax rate over the period.[18]

The main characteristics of the first stage regressions for the static panels are reported in Table 3.A5. According to the Shea partial R-square and the F-statistic, the excluded instruments are strong.

The results for various robustness tests for the static panels are provided in Table 3.A6. The upper left result is a restating of the result in regression 1 in the baseline regression in Table 3.1, and can be used as comparison for the following robustness tests.

First, alternative base year income controls are provided. The results for including 10 splines or a third degree polynomial are very similar. The result for including only a linear control is somewhat smaller, but assumingly not a sufficient control for mean reversion and income drifts (as the result becomes highly dependent on sample selections).

[18] Note that since I include year dummies in the regression analysis, what matters for estimating the elasticity of labor income is the relative differences in growth for wage earners which were more or less exposed to tax rate changes. The period t income controls make sure that the non-tax related correlation between initial income and wage growth for instance due to mean reversion is taken into account. The inclusion of base year income also allows for possible trends towards for instance increased inequality.

In the empirical literature the regressions are often weighted, with the argument that individuals contribute to the aggregate elasticity in proportion to their incomes (Saez et al., 2012). I therefore also present the results now weighted by labor income in period *t*. The obtained elasticities are somewhat larger, which could suggest that wage earners with higher income levels are more responsive. The weighted results differ somewhat more to whether splines or polynomials are used to control for mean reversion.

Then, alternative cut-off percentiles for the base year income are chosen. The baseline was to include all wage earner in percentile 33 or above. The results are insensitive to whether the lower cut-off instead is set at percentile 25 or percentile 40. This is promising, as this range of percentiles mainly serves as a control group with no or small changes in instrumented tax rate. The exclusion of the highest percentiles slightly reduce the elasticities. This again suggests that the high income earners are somewhat more responsive.

Next, I have divided the sample into two sub-periods to look at whether some periods are driving the results. I will not further divide into sub-periods as the distinction between base year income effect and tax effect get more blurred. The two sub-samples suggest that the elasticity is relatively stable over the two periods, and that for the period 1995-2004 which had relatively smaller changes in net-of-tax rate, I measure a slightly smaller elasticity.

Thereafter, I look at the impact of individuals with extreme positive or negative growth levels. The results are reasonably robust to the exclusion of 0.1 percent extreme growth files, which suggest that not only a few individuals are driving the results. The estimate drop slightly when excluding 1 percent. This means that high positive or negative growth have some impact on the results. Both this result and the result that individuals respond less in the period with small tax changes are in line with arguments from Chetty (2009b, 2012) that individuals for instance have fixed job shifting cost, such that for most individuals the change in tax rate is not large enough, to alter behavior, whereas individuals with low fixed cost or just at the threshold of making a different decision might shift job or increase effort as a consequence of the tax changes.

Next, a regression for males and females are separately estimated. The elasticity for females is lower, which is surprising as one typically finds that women are more responsive to tax changes. However, since the individuals affected by the tax changes have high income levels and probably work at least full time, it is less reason for expecting higher elasticities among women. In fact, a certain share of women in surtax position might be primary and not secondary earner in the household and therefore expected to respond equally to men. Another reason might be that very high income earners respond more (as seen above) for which there is a smaller share of women.

The main characteristics of the first stage regressions for the dynamic panels are reported in Table 3.A7. Shea partial square are somewhat lower than for the static model.

Robustness tests for the dynamic panel (according to regression 1) are summarized in Table 3.A8. In the upper left corner the reference results are restated. The reported robustness tests are chosen accordingly to in the static model with alternative lower and higher cut-off values, sub-periods, extreme growth filers and separately for gender. Again the results are quite robust to the different specification. However, the AR(2) statistic uncover that many of the specifications do not seem to fulfill the requirement that the level error terms are not serially correlated.

Table 3.A1: Summary statistics I

	1995-1999		2000-2004		2005-2008	
Variable	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Labor Income	250,285	(135,050)	318,312	(176,692)	395,421	(252,503)
Capital Income	4,675	(15,442)	6,841	(24,204)	9,255	(31,968)
Transfers (non-taxed)	9,562	(15,761)	10,932	(19,804)	10,098	(19,113)
Spouse's Net Income	177,649	(220,207)	235,937	(516,684)	294,255	(573,212)
Virtual income	228,902	(219,925)	301,149	(517,247)	367,826	(574,415)
Wealth	125,733	(888,451)	173,174	(949,424)	237,708	(2,126,603)
Marginal Tax Rate	.410	(.0832)	.408	(.0847)	.388	(.0726)

Table 3.A2: Summary statistics II

	1995-1999		2000-2004		2005-2008	
Variable	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Married	0.60	(.49)	0.55	(.5)	0.52	(.5)
Male	0.54	(.5)	0.52	(.5)	0.52	(.5)
Age	41.09	(9.9)	41.80	(10)	42.54	(10)
No. newborn children	0.17	(.42)	0.19	(.46)	0.17	(.44)
No. children under 6	0.32	(.63)	0.33	(.66)	0.31	(.64)
No. children above 6	0.55	(.86)	0.56	(.88)	0.58	(.89)
Non-western origin	0.02	(.15)	0.03	(.18)	0.04	(.19)
Residence in large city	0.12	(.33)	0.12	(.33)	0.12	(.33)
Years of education	12.23	(2.5)	12.35	(2.5)	12.50	(2.6)
Field of education (dummies)						
General subjects	0.25	(.44)	0.21	(.41)	0.28	(.45)
Humanities and arts	0.04	(.19)	0.05	(.21)	0.05	(.21)
Teacher training and pedagogy	0.07	(.26)	0.07	(.26)	0.07	(.26)
Social sciences and law	0.02	(.14)	0.03	(.16)	0.03	(.17)
Business and administration	0.16	(.36)	0.16	(.37)	0.14	(.34)
Natural sciences, vocat., techn.	0.27	(.44)	0.28	(.45)	0.25	(.43)
Health, welfare and sport	0.11	(.31)	0.13	(.34)	0.13	(.34)
Primary industries	0.02	(.13)	0.02	(.13)	0.01	(.12)
Transport and comm., safety	0.05	(.23)	0.04	(.21)	0.04	(.19)

Table 3.A3: Three-year differences

	1995-1998	1996-1999	1997-2000	1998-2001	1999-2002	2000-2003
△log(labor income)	0.149	0.153	0.144	0.132	0.129	0.123
	(0.203)	(0.211)	(0.213)	(0.217)	(0.223)	(0.217)
△log(net-of-tax)	-0.001	0.003	-0.003	-0.004	0.013	0.025
	(0.084)	(0.091)	(0.091)	(0.093)	(0.103)	(0.102)
△log(net-of-tax) instr.	0.002	0.006	-0.000	-0.002	0.020	0.035
	(0.027)	(0.049)	(0.045)	(0.037)	(0.062)	(0.077)
△log(virtual income)	0.152	0.169	0.195	0.177	0.149	0.103
	(0.312)	(0.325)	(0.333)	(0.329)	(0.353)	(0.355)
△log(virtual income) instr.	0.132	0.125	0.133	0.128	0.096	0.067
	(0.060)	(0.112)	(0.098)	(0.078)	(0.137)	(0.173)
(Capital) income shifting control	0.000	0.000	0.000	-0.618	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.226)	(0.000)	(0.000)

	2001-2004	2002-2005	2003-2006	2004-2007	2005-2008
△log(labor income)	0.103	0.091	0.109	0.146	0.177
	(0.216)	(0.215)	(0.218)	(0.228)	(0.225)
△log(net-of-tax)	0.031	0.039	0.052	0.034	0.003
	(0.102)	(0.093)	(0.088)	(0.090)	(0.084)
△log(net-of-tax) instr.	0.040	0.044	0.057	0.038	0.008
	(0.082)	(0.066)	(0.057)	(0.047)	(0.052)
△log(virtual income)	0.046	0.058	0.048	0.172	0.227
	(0.367)	(0.369)	(0.380)	(0.361)	(0.361)
△log(virtual income) instr.	0.045	0.038	0.029	0.089	0.148
	(0.180)	(0.151)	(0.123)	(0.090)	(0.110)
(Capital) income shifting control	0.567	0.000	-1.577	-1.828	-1.827
	(0.238)	(0.000)	(0.705)	(0.788)	(0.765)

Table 3.A4: Change in net-of-tax rate and labor income growth by percentile groups

	1995-1998	1996-1999	1997-2000	1998-2001	1999-2002	2000-2003
Percentile 33-54						
△log(net-of-tax) instr.	0.01	0.04	0.03	0.01	0.01	0.02
△log(labor income)	0.16	0.16	0.15	0.14	0.14	0.13
Percentile 55-74						
△log(net-of-tax) instr.	-0.00	-0.02	-0.02	-0.02	0.06	0.09
△log(labor income)	0.15	0.15	0.15	0.13	0.13	0.13
Percentile 75-94						
△log(net-of-tax) instr.	-0.00	0.00	0.00	0.00	0.00	-0.00
△log(labor income)	0.14	0.15	0.13	0.12	0.12	0.11
Percentile 95-99						
△log(net-of-tax) instr.	-0.00	0.00	0.00	-0.00	-0.00	0.00
△log(labor income)	0.14	0.14	0.13	0.11	0.11	0.09
Percentile 100						
△log(net-of-tax) instr.	-0.00	0.00	-0.12	-0.12	-0.12	-0.00
△log(labor income)	0.09	0.09	0.06	0.05	0.02	-0.02

	2001-2004	2002-2005	2003-2006	2004-2007	2005-2008
Percentile 33-54					
△log(net-of-tax) instr.	0.03	0.00	0.00	-0.00	-0.00
△log(labor income)	0.12	0.10	0.11	0.15	0.18
Percentile 55-74					
△log(net-of-tax) instr.	0.10	0.11	0.09	0.03	-0.03
△log(labor income)	0.11	0.09	0.11	0.14	0.17
Percentile 75-94					
△log(net-of-tax) instr.	-0.00	0.03	0.08	0.08	0.05
△log(labor income)	0.09	0.08	0.10	0.14	0.18
Percentile 95-99					
△log(net-of-tax) instr.	0.00	0.02	0.06	0.04	0.02
△log(labor income)	0.06	0.06	0.09	0.14	0.16
Percentile 100					
△log(net-of-tax) instr.	0.00	0.09	0.16	0.15	0.07
△log(labor income)	-0.01	0.01	0.06	0.11	0.11

Table 3.A5: Three-year static panels: First stage regressions

	Reg 1	Reg 2	
	$\log(\frac{1-\tau_{t+3}(q_{t+3})}{1-\tau_t(q_t)})$	$\log(\frac{1-\tau_{t+3}(q_{t+3})}{1-\tau_t(q_t)})$	$\log(\frac{V_{t+3}(I_{it+3},q_{it+3})}{V_t(I_{it},q_{it})})$
$\log(\frac{1-\widehat{\tau_{t+3}}(q_t)}{1-\tau_t(q_t)})$	0.6197***	0.6386***	-0.4257***
	(0.0005)	(0.0008)	(0.0030)
$\log(\frac{\widehat{V_{t+3}}(I_{it},q_{it})}{V_t(I_{it},q_{it})})$		0.0136***	0.4288***
		(0.0003)	(0.0014)
$\log(I_{it})$		-0.0005	-0.1658***
		(0.0001)	(0.0004)
Shea partial R-square	0.156	0.093	0.039
F-statistic	1,662,667	491,807	184,662

Standard errors in parentheses. All first stage regressions include the complete set of controls (as in Table 3.1) although not reported.

Table 3.A6: Robustness tests static panel

Reference						
Net-of-tax elasticity	0.0630***	(0.0021)				
Observations	9,027,613					
Alternative mean reversion controls	Linear		10 Splines		Polynomials	
Net-of-tax elasticity	0.0407***	(0.0019)	0.0635***	(0.0021)	0.0630***	(0.0021)
Observations	9,027,613		9,027,613		9,027,613	
Weighted by labor income period *t*	Linear		10 Splines		Polynomials	
Net-of-tax elasticity	0.0725***	(0.0018)	0.1044***	(0.0020)	0.0869***	(0.0019)
Observations	9,027,613		9,027,613		9,027,613	
Alternative lower income cut-offs	Percentile 25-100		Percentile 33-100		Percentile 40-100	
Net-of-tax elasticity	0.0649***	(0.0021)	0.0630***	(0.0021)	0.0616***	(0.0020)
Observations	10,038,293		9,027,613		8,124,892	
Alternative upper income cut-offs	Percentile 33-99		Percentile 33-95			
Net-of-tax elasticity	0.0554***	(0.0022)	0.0491***	(0.0021)		
Observations	8,908,944		8,384,560			
Sub-periods	1995-2004		2000-2008			
Net-of-tax elasticity	0.0495***	(0.0031)	0.0578***	(0.0022)		
Observations	5,517,069		5,136,427			
Exclude extreme growth filers	0.1% excluded		1% excluded			
Net-of-tax elasticity	0.0617***	(0.0019)	0.0509***	(0.0016)		
Observations	9,022,758		8,966,055			
Gender separately	Female		Male			
Net-of-tax elasticity	0.0310***	(0.0035)	0.0613***	(0.0026)		
Observations	3,074,406		5,953,207			

Standard errors in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.A7: Dynamic panels: First stage regressions

	Reg 1		
	$\log\left(\frac{(1-\tau_{t+1})/(1-\tau_t)}{(1-\tau_t)/(1-\tau_{t-1})}\right)$	$\log\left(\frac{1-\tau_t}{1-\tau_{t-1}}\right)$	$\log\left(\frac{q_t}{q_{t-1}}\right)$
$\log\left(\frac{1-\widehat{\tau_{t+1}}(q_{t-1})/1-\widehat{\tau}_t(q_{t-1})}{1-\widehat{\tau}_t(q_{t-1})/1-\tau_{t-1}(q_{t-1})}\right)$	0.0261***	0.3099***	-0.0038**
	(0.0011)	(0.0006)	(0.0013)
$\log\left(\frac{1-\widehat{\tau}_t(q_{t-1})}{1-\tau_{t-1}(q_{t-1})}\right)$	-0.6353***	0.9733***	0.0165***
	(0.0015)	(0.0009)	(0.0018)
$\log(q_{t-1})$	-0.0124***	0.0277***	-0.0521***
	(0.0002)	(0.0001)	(0.0019)
Shea partial R-square	0.024	0.086	0.011
F-statistic	130,174	487,474	28,563

	Reg 2				
	$\log\left(\frac{(1-\tau_{t+1})/(1-\tau_t)}{(1-\tau_t)/(1-\tau_{t-1})}\right)$	$\log\left(\frac{1-\tau_t}{1-\tau_{t-1}}\right)$	$\log\left(\frac{V_{t+1}/V_t}{V_t/V_{t-1}}\right)$	$\log\left(\frac{V_t}{V_{t-1}}\right)$	$\log\left(\frac{q_t}{q_{t-1}}\right)$
$\log\left(\frac{1-\widehat{\tau_{t+1}}(q_{t-1})/1-\widehat{\tau}_t(q_{t-1})}{1-\widehat{\tau}_t(q_{t-1})/1-\tau_{t-1}(q_{t-1})}\right)$	0.0178***	0.3167***	1.3239***	-0.4937***	-0.0070***
	(0.0011)	(0.0007)	(0.0060)	(0.0036)	(0.0014)
$\log\left(\frac{1-\widehat{\tau}_t(q_{t-1})}{1-\tau_{t-1}(q_{t-1})}\right)$	-0.6422***	0.9826***	1.2287***	-0.3036***	0.0099***
	(0.0016)	(0.0009)	(0.0084)	(0.0050)	(0.0187)
$\log\left(\frac{\widehat{V_{t+1}}(q_{t-1},I_{t-1})/\widehat{V}_t(q_{t-1},I_{t-1})}{\widehat{V}_t(q_{t-1},I_{t-1})/V_{t-1}(q_{t-1},I_{t-1})}\right)$	-0.0048***	0.0039***	0.8776***	0.0064***	-0.0016***
	(0.0002)	(0.0001)	(0.0011)	(0.0006)	(0.0002)
$\log(\frac{\widehat{V}_t(q_{t-1},I_{t-1})}{V_{t-1}(q_{t-1},I_{t-1})})$	0.0039***	0.0049***	-0.0258***	0.8925***	-0.0026***
	(0.0003)	(0.0002)	(0.0016)	(0.0010)	(0.0004)
$\log(I_{it-1})$	0.0002	-0.0011***	0.0580***	-0.1816***	0.0008***
	(0.0001)	(0.0001)	(0.0006)	(0.0004)	(0.0001)
$\log(q_{t-1})$	-0.0123***	0.0278***	0.0167***	-0.0440***	-0.0533***
	(0.0002)	(0.0001)	(0.0008)	(0.0005)	(0.0002)
Shea Partial R-square	0.025	0.090	0.188	0.242	0.012
F-statistic	65,517	244,739	262,715	373,575	14,655

Table 3.A8: Robustness tests dynamic panel

Reference (IV-2SLS)		
Short run net-of-tax elasticity	0.0661***	(0.0061)
Long run net-of-tax elasticity	0.1434***	(0.0104)
AR(2) test statistic	-2.72	
Observations	8,572,918	

Alternative lower income cut-offs	Percentile 25-100		Percentile 40-100	
Short run net-of-tax elasticity	0.0630***	(0.0061)	0.0685***	(0.0062)
Long run net-of-tax elasticity	0.1412***	(0.0100)	0.1463***	(0.0110)
AR(2) test statistic	-13.32		2.11	
Observations	9,500,377		7,730,707	

Alternative upper income cut-offs	Percentile 33-99		Percentile 33-95	
Short run net-of-tax elasticity	0.0690***	(0.0064)	0.0652***	(0.0065)
Long run net-of-tax elasticity	0.1697***	(0.0099)	0.1738***	(0.0099)
AR(2) test statistic	-10.77		-18.70	
Observations	8,465,105		7,969,894	

Sub-periods	1995-2004		2002-2008	
Short run net-of-tax elasticity	0.0662***	(0.0081)	0.0644***	(0.0074)
Long run net-of-tax elasticity	0.1428***	(0.0143)	0.1576***	(0.0117)
AR(2) test statistic	1.36		-3.09	
Observations	5,241,192		5,662,506	

Exclude extreme growth filers	0.1% excluded		1% excluded	
Short run net-of-tax elasticity	0.0631***	(0.0058)	0.0564***	(0.0052)
Long run net-of-tax elasticity	0.1341***	(0.0101)	0.1103***	(0.0095)
AR(2) test statistic	-5.64		-11.74	
Observations	8,568,771		8,510,138	

Gender separately	Female		Male	
Short run net-of-tax elasticity	0.0491***	(0.0102)	0.0726***	(0.0078)
Long run net-of-tax elasticity	0.0970***	(0.0163)	0.1382***	(0.0139)
AR(2) test statistic	-25.69		9.67	
Observations	2,992,351		5,580,567	

Standard errors in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Chapter 4

Tax Response Inertia in Labor Supply: Effects of State Dependence in Preferences and Opportunities[1]

4.1 Introduction

In the labor supply literature there is a growing awareness of that people tend to stick to their original choice of work. Chetty, Friedman, Olsen, and Pistaferri (2011) suggest that quasi-experimental studies on microdata might fail to estimate the underlying structural elasticity due to optimization frictions and highlight the importance of incorporating adjustment costs in a structural framework. In this study, we follow this idea by developing an intertemporal structural discrete choice model, similar to e.g. Haan (2010), in which state dependence can be seen as a source of adjustment costs in labor supply.

State dependence is broadly defined as the causal effect of experiencing an event on preferences, prices or constraints relevant to future choices (Heckman, 1981). The idea in our framework is that the previous choice of working hours influence both current preferences for leisure and consumption, in addition to current job opportunities. People's subjective perception of utility for alternative choices might be dependent on their experience with previous choices, for instance due to that social costs and benefits associated with a particular working time arrangement are altered. Moreover, some individuals might choose not to participate in the labor market due to lack of job opportunities, which again improves once a beneficial shock triggers a step into employment. Our main interest is that the occurrence of positive state dependence induces a gradual adjustment in working hours to changes in the tax schedule, partly due to the fact that individuals conduct several steps before the new optimal choice

[1] This chapter is based on joint work with Zhiyang Jia, Statistics Norway.

is reached, in which the utility evaluation and job opportunities gradually change, and partly because idiosyncratic shocks vary over time such that in each period a fraction of individuals absorb into the new long run equilibrium.

Studies analyzing annual labor supply transitions have long recognized the fact that individuals' labor supply choices are highly persistent over time (see e.g. Heckman, 1981). The consensus in the literature is that the correlation over time may either be due to persistence in individual heterogeneity or due to true state dependence.[2] Persistence in an individual's propensity to choose a particular labor supply state through unobserved heterogeneity might lead to so-called spurious state dependence as labor supply choices become correlated over time.

Empirical studies are not conclusive on the issue of the importance of state dependence (with a causal interpretation) in labor supply decisions. Positive state dependence at the extensive margin of female labor supply is found using US data in Hyslop (1999) and through endogenous work experience in a life cycle framework in Eckstein and Wolpin (1989). Okamura and Islam (2009), using Japanese data, find significant evidence of true state dependence in participation for married women for both the choice of regular and non-regular work. At the intensive margin, by analyzing hours worked in the framework of a dynamic Tobit model, Cai (2010) reports no evidence of true state dependence for married Australian women after unobserved heterogeneity and serial correlation of transitory errors are controlled for. Prowse (2012) and Haan (2010), on the other hand, find significant state dependence both at the intensive and extensive margin using discrete choice models estimated on respectively British and German data.

The distinction between state dependence and persistence in observed and unobserved heterogeneity is crucial when studying the time dimension of labor supply responses to policy interventions. If state dependence can be neglected, we can expect an immediate response both to permanent and transitory policy changes, where the response from a transitory policy change disappears as soon as the policy intervention has ended. Conversely, the stronger that (positive) state dependence is present, the slower we expect the responses of a permanent policy change to accumulate over time, and the slower individuals are expected to return to their normal path after a transitory change.

We depart from a static discrete choice microsimulation model which serves as a practical tool to analyze labor supply responses to alternative counterfactual policy changes; see e.g.

[2] A third possibility is serial correlation in the error terms. We show that our implementation of individual heterogeneity can be seen as a special case of serial correlation.

Dagsvik and Strøm (2006) and Dagsvik and Jia (2012). The extensions of the model to an intertemporal setting[3] are similar in spirit to contributions by Haan (2010) and Prowse (2012), although the underlying theoretical job choice framework differs slightly. One advantage of our theoretical framework is that it allows us to distinguish between state dependence in preferences and state dependence in the set of job opportunities.

Our intertemporal modeling framework should not be confused with a so-called life cycle model. A life cycle framework, advocated for instance by Keane (2011), builds on the assumption that individuals optimize the expected utility over the remaining life cycle each period, and therefore take explicitly future preferences into consideration in the current period's choice. Life cycle models can be solved by dynamic programming, but they get highly complicated when one wishes to analyze both the extensive and intensive margin of labor supply. In contrast, our model set-up is myopic in the sense that individuals optimize for each period separately. The reason why we still denote our model as intertemporal, is that previous behavior alters the current evaluation of consumption and leisure, as well as job opportunities, although this was not explicitly taken into consideration by the individual beforehand.

We estimate the model on Norwegian administrative panel data for the subset of female wage earners in couples, although the framework in general can be used for all wage earners. By simulation exercises, we predict how labor supply responses adjust over time to various changes in the wage rate, non-labor income and the tax schedule. We find that about half the effect of a permanent policy change can be expected within the first year of the policy change, whereas 90 percent of the full behavioral response is reached after about 4 years at both the intensive and extensive margin.

This chapter is organized as follows. In Section 4.2 we present our modeling framework which we denote the intertemporal job choice model. In Section 4.3, we describe the data sources and organization of the data, followed by the empirical implementation in Section 4.4. In Section 4.5 we present the estimation results of the model including wage and income elasticities, and examples of simulation exercises. Section 4.6 concludes the paper.

[3] See Dagsvik (1994) for a theoretical consideration on extending the independence of irrelevant alternatives (IIA) assumption to an intertemporal setting

4.2 The Intertemporal Job Choice Model

In this section, we start with a brief review of the static job choice model[4] followed by a discussion of the extensions necessary to analyze labor supply behavior in an intertemporal setting.

The static job choice model (applied by e.g. Aaberge, Dagsvik, and Strøm, 1995; Dagsvik and Strøm, 2006) builds on the standard discrete choice labor supply model with a conditional logit framework (see McFadden, 1974), in which individuals are assumed to choose among feasible combinations of leisure and disposable income; see for example Soest (1995) and Creedy and Kalb (2005). The discrete choice model has practical advantages over the more traditional continuous labor supply models with marginal optimization, as the discrete models easily can deal with possibly non-linear and non-convex budget constraints (Blundell and MaCurdy, 1999). The discrete choice model is therefore a widely used tool for microsimulation. The method of discretizing is typically seen as an approximation of the traditional continuous labor supply model. However, neither the continuous nor the discrete model are able to explain the peaks in full-time and part-time hours observed in the distribution of working hours. Many authors, such as Soest (1995), apply an ad hoc adjustment to account for this observation, by introducing working hour specific (dis)utility elements.

The job choice model offers, on the other hand, a theoretical rationale for discrete choice and introduces a concept of job choice opportunities to adjust for institutional peaks in the distribution of working hours; see e.g. Dagsvik et al. (2013); Dagsvik and Jia (2012). In this framework, labor supply behavior is viewed as an outcome of agents choosing among a set of job "packages", with additional constraints on the set of available jobs. The model leads to similar practical implementation as the standard discrete choice model and allows the researchers to accommodate latent choice restrictions in a convenient way.

The job choice model is specified as follows. At any given time (period t), individuals are assumed to have preferences over a set of "jobs" where each job (indexed by z) is characterized by disposable income $C_t(z)$, hours of work $h_t(z)$, and other non-pecuniary job attributes, such as the nature of the job-specific tasks to be performed, location of the workplace etc. We assume that the set of available jobs to choose from can vary both over time and among individuals. The non-market alternative is represented by $z = 0$, with hours of work $h_t(0) = 0$. Disposable income for a given job is defined as $C_t(z) = f_t(h_t(z)w_t(z), I_t)$, where $w_t(z)$ is the

[4] The static job choice model was more briefly introduced in Chapter 2.

offered wage rate for the given job z, I_t is the non-labor income and $f_t(.)$ is the net-of-tax function. In the present study, we assume that the offered wage rate $w_t(z)$ is the same for a given individual across different jobs. The individual's utility of choosing job z at period t is represented as $U_t(C,h,z)$.

In a modeling context where "job" is the decision variable, it is necessary to specify the individual-specific set of available jobs. The difficulty is that we seldom have detailed information about jobs and choice sets. However, in practice, we are not primarily interested in the job choice per se but rather in the wage and hours of work combination that follows from the job choice. In this case, Dagsvik and Jia (2012) show that when the utility function is additively separable, namely $U_t(C,h,z) = v_t(C,h) + \varepsilon_t(z)$, the labor supply probabilities can be specified in absence of detailed information about the choice sets. They introduce a variable which represents the number of available jobs for a given hours of work h: $m_t(h)$. It can be shown that under suitable distributional assumption on the error terms $\{\varepsilon_t(z)\}$, the probability that the worker chooses a particular job with hours of work h at period t, $\varphi_t(h)$ can be written as

$$\varphi_t(h) = \frac{exp(v_t(C,h))m_t(h)}{\sum_h exp(v_t(C,h))m_t(h)} \tag{4.1}$$

This expression is analogous to a multinomial logit model with representative utility $v_t(C,h)$, weighted by the number of available jobs $m_t(h)$. Since $m_t(h)$ is not observable, one need to estimate it simultaneously with the systematic part of the utility function $v_t(C,h)$. Note that without loss of generality, we can always normalize the number of non-working opportunities to one, i.e. $m_t(0) = 1$.

In the following, we extend the job choice model to study individual's labor supply behavior over time $(t = 1,...,T)$. Although individuals are still assumed to be myopic, and therefore optimize on a period to period basis, we allow for correlation in individuals choices over time due to persistence in unobserved heterogeneity and due to true state dependence, similar to approaches by e.g. Hyslop (1999) and Haan (2010).

As mentioned above, true state dependence exists when past experience has a direct effect on individual's choice behavior. In the case of labor supply, this could happen if past employment affect for example either preferences, the price of labor (in terms of wage), or labor market constraints (Heckman, 1981; Prowse, 2012). Preferences can be affected when past experience influences the taste of working and consumption, leading to intertemporally non-separable utility functions. Haan (2010) argues that this may be due to for example peer effects. The price of labor can be affected, as past working experience contributes to accumulation of job-related human capital, which thus induce an increase in wages or other types of compensation. Labor market constraints can be affected through signaling and scarring effects from past employment, or as Hyslop (1999) argues due to higher fixed cost of searching, for those who are currently unemployed.

In most empirical analysis, state dependence is represented through parameters of taste for consumption or leisure only, where state dependence through human capital and labor market constraints are assumed to work indirectly. One nice feature of our job choice model is that it explicitly introduces labor market constraints through $m_t(h)$. So we can directly model the state dependence due to labor market constraints by allowing that $m_t(h) = m_t(h, h^*_{t-1})$, where h^*_{t-1} is last period's labor supply choice.

Persistence in unobserved individual heterogeneity is introduced, as taste for leisure and consumption may vary across observable identical individuals. It can be modeled by allowing for random coefficients in the utility function.

A related aspect is serial correlation in the error terms $\{\varepsilon_t(z)\}$ which arises as a result of random shocks which last longer than one period or unobserved individual characteristics that change slowly over time. Hyslop (1999) argues for instance that allowing for serial correlation in the error terms can correct for the endogeneity of non-labor income and fertility. A simple form of serial correlation can be introduced in the discrete choice model by allowing for choice specific permanent unobserved heterogeneity. Consider a simple representation of serial correlation structure where the error term consists of two parts: a permanent individual specific effect that does not change over time, and a transitory part. We will assume that the individual permanent effects for jobs with the same hours of work are constant. Namely: $\varepsilon_t(z) = a(h) + \zeta_t(z)$, where $\zeta_t(z)$ is i.i.d across time periods and jobs. If we assume that $a(h)$ is random across individuals and distributed according to a suitable distribution $f(.)$, we see immediately that the choice probability follows a standard random effect approach:[5]

$$\varphi_t(h|\vec{a}) = \frac{exp(v_t(C,h) + a(h))m_t(h)}{\sum_k exp(v_t(C,k) + a(k))m_t(k)} \tag{4.2}$$

[5] Since we typically do not observe jobs, it is difficult to handle the more general case where the individual permanent effect depends directly on jobs.

$$\varphi(h_1, h_2 \cdots h_T) = \int_{\Omega_{\vec{a}}} \prod_{t=1}^{T} \varphi_t(h_t|\vec{a}) f(\vec{a}) d\vec{a} \tag{4.3}$$

It follows that allowing for unobserved heterogeneity in the choice set $m_t(h)$ can be interpreted as a special type of serial correlation in the error terms $\{\varepsilon_t(z)\}$.

A more detailed specification of the empirical model is discussed in Section 4.4.

4.3 Data

The empirical application is based on individual data from the Norwegian Earnings Survey combined with information from a collection of administrative registers denoted Income Statistics for Persons and Families.

The Norwegian Earnings Survey (also called Wage Statistics) collects monthly data from the employer in September/October each year since 1997 (see documentation Statistics Norway, 2006, 2009). The statistics are collected from a stratified selection of Norwegian enterprises with at least 3-5 employees (depending on industry). In private sector 50-60 percent and in public sector 100 percent of the employees are covered by the statistics. In total, we have information on more than 70 percent of Norwegian wage earners. The key variables for our analysis are reported working hours and contractual monthly pay for the reference month.

Income Statistics for Persons and Families are collected from various administrative registers since 1993 and cover the complete Norwegian population. The statistics include tax return, education and family statistics (see Statistics Norway, 2005). We use yearly information on labor (wage and self-employed income) and non-labor income such as capital income and transfers, as well as various individual characteristics such as gender, age, education and origin. Spouse and children are identified through a unique household number. Personal identification numbers allow us to combine these data with the information from the earnings survey.

For this analysis, we construct a balanced panel over the period 2003-2007, where the first year is used for initial conditions only. We focus on the subgroup of women in couples aged 25-62, who can be characterized as wage earners or potential wage earners, and exclude self-employed, disabled and unemployed. Further, we assume for simplicity that husband's income is taken as given, and restrict the analysis to couples where the husband's total pre-tax income level exceed yearly NOK 150,000 (about EUR 20,000).[6]

[6] The restriction affects less than five percent of the couples. The reason for exclusion of couples with lower

The sample of women are divided into five discrete labor supply choices: No work, short part time, long part time, full time and overtime, characterized by working hours in the range of 0, 1-19, 20-34, 35-40 and more than 40 hours per week. We construct a measure for hourly wage rate by using the contractual monthly pay divided by monthly contractual working hours. Next, we estimate a Mincer wage regression allowing for selection effects for participation, according to Heckman (1979). The regression output is reported in the Appendix. In order to be consistent, and to alleviate systematic bias in the computed wage,[7] we use the predicted wage rate for all individuals.

As we prefer to include the full sample of individuals to avoid attrition/selection effects over time, we need to categorize individuals not observed in the wage statistics into the five different working hour alternatives. This is done, similarly to Pronzato (2012) by regressing the monthly wage of individuals in full time job, and using the predicted monthly wage as basis for categorizing. Individuals with yearly income about 12 times the monthly wage are for instance assumed to work full time. The cut-offs are calibrated by adjusting the simulated to the actual distribution of working hours for individuals in the wage statistics sample. Non-participants are defined as having yearly wage income less than NOK 5,000 or less than 0.3 times the predicted monthly wage rate.

Next, we simulate disposable income ("consumption") for each working time alternative for each year by adding the predicted wage income (based on the median working hours in each alternative) with capital income, transfers and husbands income. Taxes are simulated for the household according to the applicable tax schedule for each year, and deducted such that net disposable household income remains.

In the following analysis we pick a ten percent random sample of the individuals in order to reduce the computational time. The results are robust to alternative ten percent draws. The summary statistics of the restricted sample is presented in Table 4.1 and Table 4.2. Table 4.1 presents the relevant income measures over time, whereas Table 4.2 describes the characteristics included as taste modifying variables in the utility function, in addition to years of education which is included in the specification of the choice set.[8] Figure 4.1 depicts the observed share of individuals in each working time arrangement. The labor market participation

income is in order to fulfill the assumption that husbands income is considered exogenous and that the womens' choice is not constrained. Moreover, when conditioning on a minimum household income level, we can abstract from various welfare transfers.

[7] Bias in working hours lead to negatively correlated bias in the computed wage.

[8] Years of education are imputed for a small fraction of individuals with missing observations. A few individuals are coded with 0 years of education.

rate is clearly increasing over time and there seems to be a shift from part time to full time and overtime.[9] Both suggest an increase in female labor supply over time, although this pattern is somewhat exaggerated due to the balanced panel. Table 4.3 presents the transition probabilities for year 2005 to 2006.[10] The diagonal entries of the matrix are substantial, which suggests strong persistence in labor supply over time.

Table 4.1: Sample characteristics over time

	Mean income (NOK)				
Variable	2004	2005	2006	2007	N
Labor income (yearly)	247,564	260,500	277,688	299,909	29,848
Labor income (Reference month)	20,859	21,530	22,624	24,109	20,353
Non-labor income	28,357	39,364	24,822	26,569	29,848
Husbands total income	515,546	580,154	515,479	573,527	29,848

Table 4.2: Pooled sample characteristics

Variable	Mean	Std. Dev.	Min.	Max.
Age	44.2	8.9	25	62
No. children under 3	0.22	0.5	0	4
No. children under 6	0.44	0.74	0	6
No. children under 12	0.88	1.05	0	10
Years of education	12.1	2.6	0	20

[9] Over the time period of consideration a tax reform was implemented with lower marginal surtax rates (see Chapter 2 and Chapter 3) and with some smaller changes in the basic tax allowance. We have deliberately chosen to estimate the model over a period with changes in the tax schedule in order to enhance the structural estimates of the intertemporal model.

[10] The annual transition probabilities are similar for the other years.

Figure 4.1: Observed choice probabilities

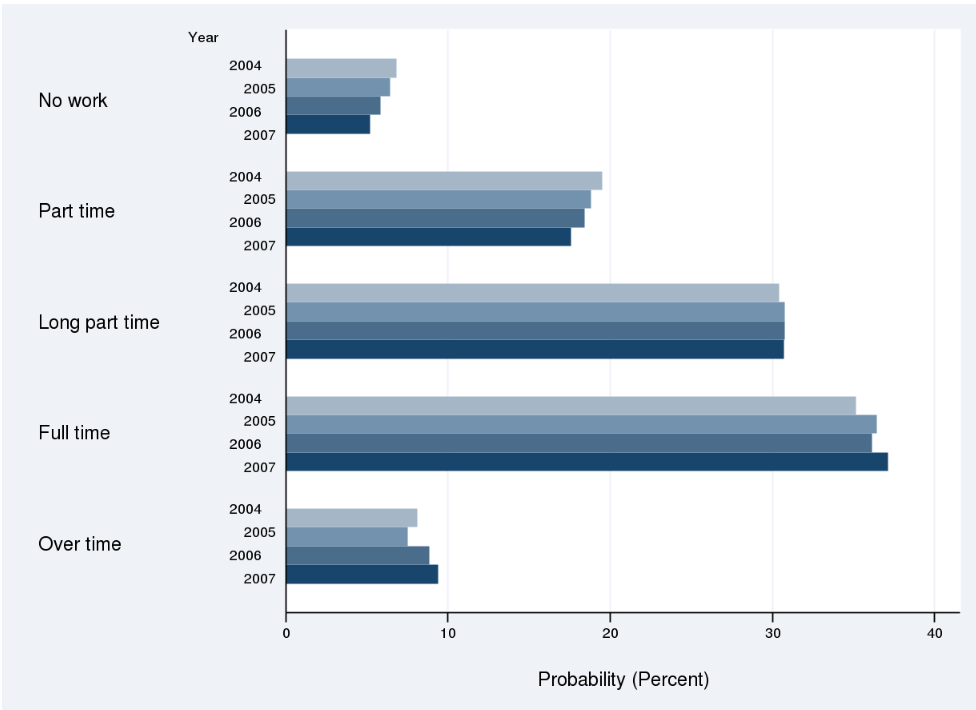


Table 4.3: Observed transition probabilities 2005-2006

		2006				
		No work	Short part time	Long part time	Full time	Over time
	No work	**81.9%**	15.8%	1.2%	0.9%	0.3%
	Short part time	2.7%	**76.0%**	16.5%	4.1%	0.7%
2005	Long part time	0.1%	7.7%	**77.3%**	12.7%	2.2%
	Full time	0.1%	1.6%	9.1%	**79.8%**	9.4%
	Over time	0.0%	1.9%	6.4%	31.1%	**60.7%**

4.4 Empirical Specification

In this section, we discuss the empirical specification of the model. We need to specify the functional form for the deterministic part of the utility function $v(C,h)$ and the opportunity measure $m(h)$, both allowing for unobserved heterogeneity and true state dependence. Following Haan (2010) and Prowse (2012) we assume that past employment history influence current period's labor supply decision, only through last period's decision. In other words, the labor market behavior follows a first order Markovian structure over time. Further, we assume that last period's decision influence both current individual preferences as well as the current available job choice set.

As regards the systematic part of the utility function, we follow Dagsvik and Strøm (2006); Dagsvik and Jia (2012) and make use of a generalized Box-Cox function in leisure and disposable income. This functional form has the advantage of being strictly concave under some simple constraints on the parameters.[11] We follow Dagsvik and Jia (2012) and assume;[12]

$$v_t(C,h) = \alpha_{0i} \cdot \frac{((C_t - C_0)^{\alpha_1})}{\alpha_1} + \beta_{0i} \cdot \frac{(L_t)^{\beta_1}}{\beta_1} \tag{4.4}$$

where leisure is defined as $L_t = (80 - h_t)/80$ with h_t equal discrete weekly working hours and $C_t - C_0 = (di - 50{,}000 \cdot \sqrt{hh})/10{,}000$ where di measure household disposable income and hh is the number of individuals in the household. The deduction of NOK 50,000 can be interpreted as some minimum substantial level of consumption. The parameters α_{0i} and β_{0i} are both allowed to differ among individuals representing permanent unobserved heterogeneity in preferences. The leisure coefficient, β_{0i}, is additionally defined to depend on observed individual characteristics, $\mathbf{x_{it}}$.

$$\beta_{0i} = b_{0i} + \mathbf{b_1'} \cdot \mathbf{x_{it}} + \mathbf{b_2'} \cdot \mathbf{L_{i,t-1}} + \mathbf{b_3'} \cdot \mathbf{L_{i,0}}$$

True state dependence in preferences is accounted for through the inclusion of last period's choice L_{t-1}, while initial period choice L_0 is included to solve the initial condition problem

[11] Many researchers apply a polynomial (in most cases, a quadratic) specification which is quite flexible and easy to deal with (linear in parameters). One problem with a polynomial functional form is that it is not globally monotone in consumption or leisure. See Dagsvik et al. (2013) for a more detailed discussion on this issue.

[12] The general Box Cox utility function does also include an interaction term of consumption and leisure. We have neglected this additional term as we did not find any significant effect.

following the method by Wooldridge (2005). A similar strategy for the structure of state dependence can be found in e.g. Haan, Kemptner, and Uhlendorff (2012).

As discussed earlier, individual's working history may have a direct effect on individual's current labor supply behavior through labor market restrictions. Since labor market restrictions are explicitly modeled in the job choice framework, it is possible to check this empirically by including working history in the opportunity measures $m(h)$. The opportunity measure $m(h)$ can be separated in the following manner:

$$m(h) = \theta g(h)$$

where $\theta = \sum_{h>0} m(h)$ and $g(h) = m(h)/\theta$. We denote $g(h)$ the opportunity distribution of hours, which can be interpreted as the fraction of jobs available to the agent with offered hours of work equal to h. The parameter θ describes the total number of jobs available to the agent.[13] Under this notation, we have

$$\varphi_t(h) = \frac{exp(v_t(C,h))g_t(h)}{v_t(C,0))\theta_t^{-1} + \sum_{h>0} exp(v_t(C,h))g_t(h)} \tag{4.5}$$

for h>0 and similarly for h=0.

The opportunity distribution of hours $g(h)$ is considered to be stemmed from the demand side due to institutional restrictions as a results of centralized negotiations between labor unions and employers' associations. We treat $g(h)$ as constant over time as institutional settings are relatively stable over time. Further, we assume that the opportunity distribution is uniformly distributed among working time alternatives, except for a possible peak (estimated within the model) for full time jobs.

$$log(g_t(h)) = \begin{cases} g & if\,working\,full\,time \\ 0 & otherwise \end{cases}$$

Regarding the size of choice set θ (relative to non-participation), we assume it depends on individual time-constant unobserved effect, education and last periods working time experience. Since labor market conditions are closely related to the general macro economic conditions, we also include time dummies (δ_t).[14] By allowing for that θ depends on last period's choice

[13] Note that θ can be extended to include fixed cost of working; see Cogan (1981).

[14] An alternative strategy would be to include macroeconomic indicators such as unemployment rates. The ob-

L_{t-1}, we explicitly model the state dependence through labor supply constraints. As before, we include L_0 to control for initial conditions.

$$log\,\theta_{it} = \mu_i + \delta_t + \gamma_1 \cdot edu_{it} + \boldsymbol{\gamma_2} \cdot \mathbf{L_{i,t-1}} + \boldsymbol{\gamma_3} \cdot \mathbf{L_{i,0}}$$

μ_i represents individual unobserved heterogeneity in the number of jobs available, and is jointly estimated with unobserved heterogeneity in preferences.

As discussed earlier, there may also be serial correlations in the error terms over time. In Section 4.2, we demonstrated that we can account for certain types of autocorrelation through random effects specifications, see equation (4.2). Here, we make an even simpler assumption in the autocorrelation structure: $\varepsilon_t(z) = a + \zeta_t(z)$ if $z > 0$, and $\varepsilon_t(z) = \zeta_t(z)$ if $z = 0$, where a is constant across different time period, but different across individuals. The assumption may be restrictive since it implies that the serial correlations among working states are identical. Using (4.2) and (4.6), we see immediately that we can separately identify a from θ_{it}, such that a random effect in μ_i accounts for this simple autocorrelation.

We assume that the random variables $\{\alpha_{0i}, b_{0i}, \mu_i\}$ follow a discrete distribution (see Heckman and Singer, 1984) and take values on k mass points with probability $\pi_k \in (0,1)$.[15] This leads essentially to a finite mixture model and has the advantage that unobserved individual characteristics can be handled flexibly without imposing a parametric structure. We estimate simultaneously the mass points and corresponding probabilities. In principle we impose no restrictions on possible correlation structure between these random variables, see e.g. Haan (2010) for similar arguments.

With the assumptions described above, the joint likelihood for individual to choose the sequence $(h_1, h_2 \cdots h_T)$ can be written as

$$\varphi(h_1, h_2 \cdots h_T) = \sum_k \pi_k \prod_t \varphi_t(h|\alpha_{0i}^k, \mu_i^k) \tag{4.6}$$

with the conditional probabilities for each mass point given by equation (4.5).

served increase in participation over time might, however, also reflect an increase in childcare facilities or be due to the characteristics of a balanced sample.

[15] Another interesting way to specify the bi-variate distribution is to use the so-called “two-factor loading” models, see e.g. Haan and Uhlendorff (2012) for an application in labor supply modeling.

4.5 Results

In the following, we present the estimated parameters of the labor supply model.[16] Based on the estimated model, we then derive the wage and income elasticities, and demonstrate how the model can be used to perform counterfactual tax reforms with the focus on the dynamics of responses. Further, we discuss in more detail the importance of state dependence relative to unobserved heterogeneity in our modeling context, and provide some robustness checks on alternative modeling of unobserved heterogeneity.

4.5.1 Model Estimates and Fit

In our main results we have chosen to include two mass points, each with an estimated vector of $\{\alpha_{0i}, b_{0i}, \mu_i\}$ and an associated estimated probability for each vector. As two mass points might be considered restrictive, we report results for a robustness test with four mass points at the end of this section.[17] Although more than two mass points slightly improves the fit of the model, we experienced problems with a small share of individuals with zero or slightly negative utility for consumption. We are reassured by the fact that the predicted wage elasticities over time barely differ between the two, three or four mass points model.

The estimated model coefficients for the two mass points case are presented in Table 4.4. The coefficients for preferences imply positive marginal utility of both consumption and leisure, and as α_1 and β_1 are less than 1, it can be shown that the utility function is concave with respect to both arguments. The taste of leisure varies with individual characteristics such as age and number of children. Utility from leisure is increasing with age (in the relevant age interval), and having small children also increase the value of leisure, especially for women with children under three years old.

For the opportunity measure (θ), which can be interpreted as a proxy for the number of available jobs, we see a strong time effect which suggest more jobs available to women from year 2004 to year 2007 (note that the negative estimates refer to the inverse of the log opportunity measure). This is in general consistent with the decline of the unemployment rate in Norway during the same period (4.5 percent in 2004 to 2.5 percent in 2007, Statistics Norway, 2013). We suspect, in addition, this may be connected to the increase in the supply of public daycare

[16] The model is estimated utilizing the ml-command in STATA.

[17] Akaike/Bayesian information criterion (AIC/BIC) can be used to decide how many mass points are sufficient to include. In empirical studies one rarely see more than three or four mass points.

over the same period.[18] In contrast to our expectations, we find no significant effect of education on women's job choice set. The estimated opportunity density shows a clear peak at full time.

The estimates for state dependence in preferences ($b_{22} - b_{25}$) are negative, which suggest a negative effect of last periods working experience on the value of leisure the current period. This effect seems to increase with the intensity of work the previous period. Regarding state dependence in the opportunity measure (choice set) we find accordingly a positive effect of last periods working experience on this periods choice set (note that negative estimates, $\gamma_{22} - \gamma_{25}$, refer to the inverse of the opportunity measure). The effect of last periods choice on the opportunity measure, in contrast to the preference parameters, does not vary much with intensity of work, but has a strong effect on participation. This is consistent with our expectations: Individuals' choices are more constrained once they leave the labor market fully, due to for example, increased search cost, discouraged worker effect etc. We find strong and significant effects for the initial conditions (initial period labor supply choices), which indicates the importance of accounting for this.

Note that our model estimates a common distribution for the random effects for all individuals within the sample (the prior distribution). However, based on the labor supply dynamics, we can refine this to obtain individual specific distributions (the posterior distributions.). This method is labeled as the empirical Bayes method in the literature; see Skrondal and Rabe-Hesketh (2004) and Haan (2010). In particular, we define the individual weights

$$w_{ik} = \frac{\pi_k \prod_t \varphi_{it}(h|\alpha_0^k, \mu^k)}{\sum_j \pi_j \prod_t \varphi_{it}(h|\alpha_0^j, \mu^j)}.$$

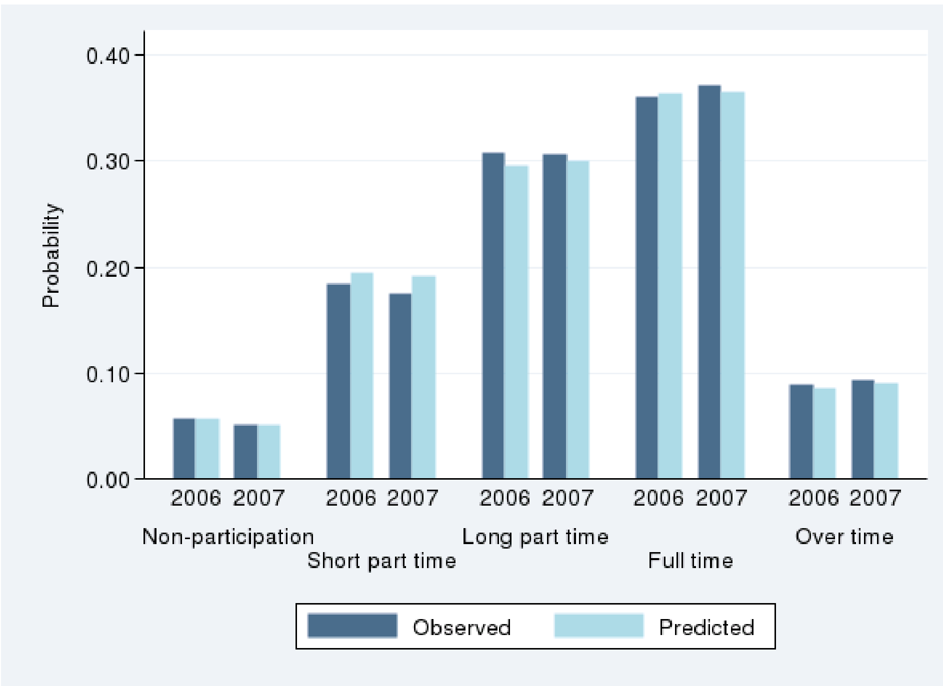
and use these weights when we predict the individuals' labor supply behavior over time.

The predicted marginal labor supply shares are very close to the observed ones, as shown in Figure 4.2. The predicted labor market transitions, as shown in Table 4.6, can be compared to the observed transition probabilities in Table 4.3. While the labor market persistence for non-participation is very well replicated, the persistence at the intensive margin seems somewhat underestimated. One reason could be a relatively restrictive modeling of state dependence. Haan (2010), in contrast, allows for that last period's labor market status have a unique effect

[18] The daycare coverage of children aged 1-5 years increased from 72.1 percent in 2004 to 84.3 percent in 2007 (Statistics Norway, 2013). The rise in coverage might also partly reflect increased demand.

for each different working hour alternative, and consequently include dummies for each possible combination. Although the fit is better, one looses some of the structural interpretation. Another reason for our slightly underestimated persistence at the intensive margin could be a too restrictive modeling of the autocorrelation in error terms (see Hyslop, 1999 or Prowse, 2012 for a less restrictive modeling of autocorrelation).[19]

Figure 4.2: Observed and predicted state probabilities, year 2006 and 2007



[19] We have not been successful in extending our model along these lines due to numerical challenges with five working time categories.

Table 4.4: Parameter estimates of the labor supply model for women in couples

		Coefficient	Std error
Consumption			
Constant (Scale 10^{-4})			
Mass point 1	α_{01}	0.7365***	(0.0227)
Mass point 2	α_{02}	7.2830***	(0.1953)
Probability distribution $(\alpha_{0i}, b_{0i}, \mu_i)$			
Probability mass point 1	$p1$	0.7583***	(0.0038)
Probability mass point 2	$p2$	0.2417***	(0.0038)
Exponent	α_1	0.7172***	(0.0070)
Leisure			
Constant (Scale $1/80$)			
Mass point 1	b_{01}	4.4857***	(0.2077)
Mass point 2	b_{02}	16.2241***	(0.4220)
Taste modifiers			
Age (Scale 10^{-1})	b_{11}	-0.1871*	(0.0744)
Age squared (Scale 10^{-2})	b_{12}	0.0362***	(0.0081)
No. children under age 3	b_{13}	0.3231***	(0.0205)
No. children under age 6	b_{14}	0.0232	(0.0161)
No. children under age 12	b_{15}	-0.0878***	(0.0099)
State dependence			
Choice 2 period t-1	b_{22}	-0.3855***	(0.1016)
Choice 3 period t-1	b_{23}	-2.3744***	(0.1071)
Choice 4 period t-1	b_{24}	-3.6148***	(0.1146)
Choice 5 period t-1	b_{25}	-4.6726***	(0.1261)
Initial conditions			
Choice 2 period 2003	b_{32}	0.5793***	(0.0613)
Choice 3 period 2003	b_{33}	0.0568	(0.0609)
Choice 4 period 2003	b_{34}	-0.3783***	(0.0614)
Choice 5 period 2003	b_{35}	-0.8517***	(0.0654)
Exponent	β_1	-2.8618***	(0.0296)

Opportunity measure (Inverse)			
Constant			
Mass point 1	μ_1	2.3300***	(0.1376)
Mass point 2	μ_2	1.3149**	(0.4271)
Education			
Years of education	γ_1	0.0092	(0.0105)
State dependence			
Choice 2 period t-1	γ_{22}	-4.0660***	(0.0808)
Choice 3 period t-1	γ_{23}	-5.6093***	(0.2485)
Choice 4 period t-1	γ_{24}	-4.7024***	(0.2896)
Choice 5 period t-1	γ_{25}	-1.8807***	(0.4382)
Initial conditions			
Choice 2 period 2003	γ_{32}	-1.4803***	(0.0753)
Choice 3 period 2003	γ_{33}	-1.4499***	(0.1377)
Choice 4 period 2003	γ_{34}	-0.9038***	(0.1778)
Choice 5 period 2003	γ_{35}	-0.4861	(0.3146)
Year dummies (Base: 2004)			
Year 2005	δ_{2005}	-0.1553*	(0.0723)
Year 2006	δ_{2006}	-0.3043***	(0.0731)
Year 2007	δ_{2007}	-0.3729***	(0.0741)
Opportunity density			
Full time peak	g_4	1.6598***	(0.0149)
Number of individuals		29,757	

Table 4.5: Predicted choice probabilities

	2004	2005	2006	2007
No work	6.7%	6.4%	5.8%	5.1%
Short part time	20.8%	20.2%	19.5%	19.3%
Long part time	27.9%	28.8%	29.5%	29.8%
Full time	35.9%	35.7%	36.6%	36.7%
Over time	8.7%	8.9%	8.5%	9.0%

Table 4.6: Predicted transition probabilities 2005-2006

		2006				
		No work	Short part time	Long part time	Full time	Over time
2005	No work	**81.9%**	14.3%	3.3%	0.5%	0.0%
	Short part time	2.8%	**71.3%**	22.2%	3.7%	0.0%
	Long part time	0.0%	12.3%	**66.7%**	20.8%	0.3%
	Full time	0.0%	3.8%	12.3%	**71.6%**	12.4%
	Over time	0.0%	0.5%	2.2%	44.8%	**52.3%**

4.5.2 Wage and Income Elasticities

Wage and income elasticities predicted from the model are presented in Table 4.7 and 4.8. Further, Figure 4.3 provides a graphical illustration of the wage elasticities over time. The wage elasticities are simulated by increasing the wage rate by ten percent and measuring the percentage change in predicted working hours for each individual over time; the average elasticity over individuals is reported. Non-labor income elasticities are simulated by increasing the non-labor income (here husband's gross income) by ten percent and measuring the percentage change in predicted working hours. For each individual we compare a reference path with actual 2007-wage income to an alternative path with a transitory or permanent increase in the wage rate, both paths with the 2007 tax system and non-labor income left unchanged. This is conducted by starting with the predicted probability distribution for 2004 by the observed initial conditions and obtaining a probability distribution for the next year according to the following formula $\varphi_{it}(h_t|h_0) = \sum_{h_{t-1=j}} \varphi_{it}(h|h_{t-1},h_0) \cdot \varphi_{it-1}(h_{t-1}|h_0)$.

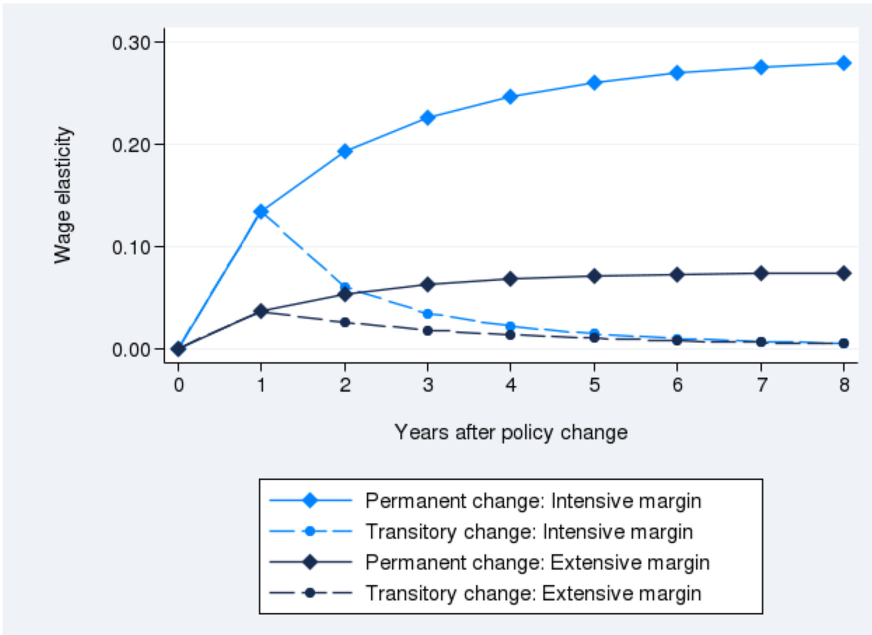
Table 4.7: Predicted wage elasticities over time (ref. year 2007)

	Wage elasticities (Std error)			
	Transitory change		Permanent change	
	Intensive margin	Extensive margin	Intensive margin	Extensive margin
1st year	0.13 (0.0010)	0.04 (0.0006)	0.13 (0.0010)	0.04 (0.0006)
2nd year	0.06 (0.0004)	0.03 (0.0004)	0.19 (0.0013)	0.05 (0.0009)
3rd year	0.03 (0.0002)	0.02 (0.0003)	0.23 (0.0015)	0.06 (0.0011)
4th year	0.02 (0.0002)	0.01 (0.0003)	0.25 (0.0015)	0.07 (0.0013)
5th year	0.01 (0.0001)	0.01 (0.0002)	0.26 (0.0016)	0.07 (0.0015)
10th year	0.00 (0.0000)	0.00 (0.0001)	0.28 (0.0017)	0.07 (0.0019)

The results suggest that within the first year of a permanent change in the wage rate about 50 percent of the full labor supply response can be expected, whereas 90 percent of the response is reached after about 4 years both at the extensive and intensive margin.

In contrast to the general perception that wage elasticities are higher at the extensive margin (see e.g. Heckman, 1993), we find the opposite. This might be specific to the Norwegian economy. Along with the increased participation rate for women, one might see a general trend towards lower participation elasticities. For comparison I estimated, in Chapter 2, elasticities

Figure 4.3: Predicted wage elasticities over time



of 0.09 at the extensive margin and 0.29 at the intensive margin for women in couples in a static version of the job choice model, so interestingly the results obtained from the static model are very similar to the long run wage elasticities in our intertemporal framework.

The income elasticities are, as expected small negative.

Table 4.8: Predicted income elasticities over time (ref. year 2007)

	Income elasticities (Std error)			
	Transitory change		Permanent change	
	Intensive margin	Extensive margin	Intensive margin	Extensive margin
1st year	-0.02 (0.0005)	-0.01 (0.0006)	-0.02 (0.0005)	-0.01 (0.0006)
2nd year	-0.01 (0.0002)	-0.01 (0.0005)	-0.04 (0.0007)	-0.01 (0.0009)
3rd year	-0.01 (0.0001)	-0.00 (0.0003)	-0.04 (0.0008)	-0.01 (0.0010)
4th year	-0.00 (0.0001)	-0.00 (0.0003)	-0.05 (0.0009)	-0.02 (0.0012)
5th year	-0.00 (0.0001)	-0.00 (0.0002)	-0.05 (0.0010)	-0.02 (0.0013)
10th year	-0.00 (0.0000)	-0.00 (0.0001)	-0.05 (0.0010)	-0.02 (0.0015)

4.5.3 Tax Simulations

In the following, we demonstrate how the estimated model can be used for evaluating counterfactual policy reforms. In order to interpret the simulation results, we briefly describe the tax schedule for Norwegian wage earners in year 2007, which serves as the benchmark when conducting alternative tax policies.

Norway has a dual income tax system implemented through two separate tax bases denoted "ordinary income" and "personal income". Ordinary income consist of all taxable income from labor and capital after basic allowance and other tax deductions are subtracted, and is taxed with a flat rate of 28 percent. Personal income includes gross labor income and serves as the tax base for social security contribution (7.8 percent for wage earners) and a two tier surtax rate system. In 2007, the first surtax tier was applied at a rate of 9 percent to incomes above NOK 400,000, and the second tier of 12 percent applied to income in excess of NOK 750,000. Two separate tax classes apply to Norwegian residents. The majority are taxed individually in tax class 1 (more than 90 percent). Single parents are taxed in tax class 2, in which the "personal allowance" is doubled. Tax class 2 also applies to married couples if joint assessment is beneficiary (usually only one-income families).

We consider three counterfactual policy changes presented in Table 4.9 with 2007 as a starting point and reference path. The first policy change is the abolishment of the second tax class, which in our context means an abolishment of the possibility of joint taxation for women in couples. The second policy change is an abolishment of the basic and personal tax allowance (corresponds to maximum NOK 100,800 in tax class 1 in 2007), and in the third policy change we consider the abolishment of the two-tier surtax schedule. We briefly comment on the results for each of the three policy changes.

The first counterfactual policy change, the abolishment of joint taxation, leads to a less attractive alternative of non-participation for married women as the husband looses the additional personal allowance. As expected we see a negative predicted effect on the probability of choosing non-participation, and a correspondingly higher probability of longer part time and full time. The non-participation probability decreases with about 0.4 percentage points (from 5.1 percent to 4.7 percent) the first year and with about 0.9 percentage points in the long run. Predicted average working hours increase only slightly by 0.1-0.3 hours per week.

The second policy change, the abolishment of the basic tax allowance decreases the incentives to participate in the labor market. However, one special feature of the Norwegian basic tax allowance ("minstefradrag") is that it is increasing with earnings until a maximum level

Table 4.9: Simulation results - 3 hypothetical tax reforms

	Change in percentage points (relative to reference path)					
Abolish joint taxation	No work	Short part time	Long part time	Full time	Over time	ΔWorking hours
1st year	-0.4	+0.1	+0.3	+0.1	+0.0	0.14
2nd year	-0.6	+0.1	+0.3	+0.2	+0.0	0.17
3rd year	-0.7	+0.1	+0.3	+0.3	+0.0	0.21
10th year	-0.9	-0.1	+0.3	+0.6	+0.1	0.34
Abolish basic tax allowance	No work	Short part time	Long part time	Full time	Over time	ΔWorking hours
1st year	+0.2	+3.0	-2.9	-0.4	+0.0	-0.51
2nd year	+0.4	+3.9	-3.4	-0.9	+0.0	-0.70
3rd year	+0.5	+4.3	-3.5	-1.3	-0.0	-0.82
10th year	+0.8	+5.0	-3.5	-2.1	-0.2	-1.11
Abolish surtax schedules	No work	Short part time	Long part time	Full time	Over time	ΔWorking hours
1st year	-0.0	+0.8	-0.7	-0.3	+0.1	-0.14
2nd year	0.0	+1.1	-0.7	-0.5	+0.1	-0.17
3rd year	0.0	+1.2	-0.8	-0.6	+0.1	-0.22
10th year	+0.0	+1.4	-0.8	-0.8	+0.1	-0.27

is reached (amounting to NOK 63,800 for wage earners in 2007). So, in currency amount, only women working at least long part time or full time are fully profiting from the basic tax allowance. At the same time the tax policy affects husbands disposable income such that there is a negative income effect both from own and husband's disposable income which in itself is expected to increase the number of working hours. The results suggest an increase in non-participation and short part time and a decrease in long part time and full time. We see that all changes are rather sluggish in particular for adjustments in non-participation and full time.

The last counterfactual policy change we consider, is an abolishment of the surtax schedule, leaving the tax schedule flat, apart from the features of the basic tax allowances and deductions. As expected, this induce an increase in overtime work, since a considerable proportion of the women would be in surtax rate position conducting a job characterized with overtime work. Depending on the wage rate, the choice of working full time would, for a certain fraction of the women, also be affected by the surtax rate abolishment. This effect of increased working hours seems surprisingly low. Instead, it seems like the effect through husband's income is stronger. The majority of husbands in our sample are affected by the abolishment of the surtax rates, which leads to a positive income effect for the women in our modeling framework. The model predicts an overall negative effect on the probability of working full time and long part time. Consequently there is a negative overall effect on working hours. The model predicts somewhat faster responses, where more than half the long run effect is reached already the first year.

Similar analysis can be conducted for a range of counterfactual policy changes, these three analysis serve primarily as examples of how the estimated model can be used for policy makers in practice.

4.5.4 True State Dependence and Unobserved Heterogeneity

One of the crucial determinants for measuring the time dimension of responses is as mentioned to separate between persistence in unobserved heterogeneity and true state dependence. In this subsection we treat this issue in more detail, by showing results from some alternative specifications of the model, now using a three percent sample (the main analysis was based on a ten percent sample) in order to further limit computational time.

In Figure 4.4 and Table 4.10 we summarize the fit and wage elasticities over time for alternative specifications. The reference case equates to the main specification described in the previous sections, apart from the smaller sample size. The second model is specified without any unobserved heterogeneity in preferences and the choice set, and the third model is specified without state dependence.[20]

As we can see from Figure 4.4, the persistence is better captured by the reference case than for the model in which we did not allow for unobserved heterogeneity. There is in particular a large difference for the probability of staying in short or long part time, indicating that unobserved heterogeneity should not be neglected. When not allowing for state dependence the predicted persistence decrease further. This might suggest that state dependence is accounting for most

[20] We include initial conditions although state dependence is excluded

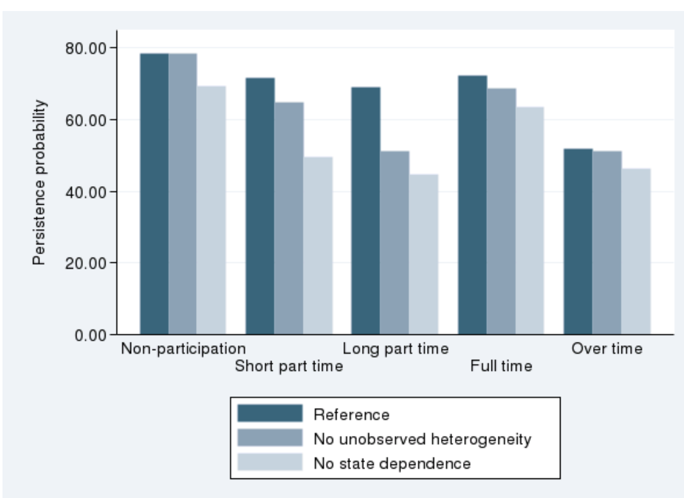
of the observed persistence in individuals' choices over time.[21]

Table 4.10: Wage elasticities

	Wage elasticities					
	Intensive margin			Extensive margin		
	Short-run[a]	Long-run[b]	Time[c]	Short-run[a]	Long-run[b]	Time[c]
Reference	0.13	0.28	7	0.04	0.08	4
No unobserved heterogeneity	0.18	0.51	6	0.08	0.21	10
No state dependence	0.21	0.21	0	0.16	0.16	0

a First year responses, *b* 10th year responses, *c* Years until 95% of long run effect reached

Figure 4.4: Predicted persistence probabilities for the transition year 2006-2007



[21] The difference between the reference model and the model without unobserved heterogeneity is slightly larger when modeling unobserved heterogeneity with four mass points.

4.5.5 Robustness Check: Unobserved Heterogeneity

In Table 4.11 we summarize the wage elasticities over time for a range of alternative assumption on unobserved heterogeneity. We first show the results for a specification with four mass points instead of two as described in the beginning of this section. Then we allow for random effects in wages by adding 30 normal distributed draws.[22] In the last alternative we model the random effects as normal distributed in which α_{0i} and β_{0i} are jointly randomly determined, whereas μ_i follows from an independent draw.[23]

Table 4.11: Alternative specifications: Wage elasticities

	Wage elasticities					
	Intensive margin			Extensive margin		
	Short-run^a	Long-run^b	Time^c	Short-run^a	Long-run^b	Time^c
Reference (2 mass points)	0.13	0.28	7	0.04	0.08	4
4 mass points in $(\alpha_{0i}, b_{oi}, \mu_i)$	0.15	0.26	6	0.05	0.09	6
Add normal distr. wage random effects	0.15	0.31	7	0.04	0.08	5
Normal distributed (α_{0i}/b_{oi}) and μ_i	0.18	0.33	5	0.09	0.13	5

a First year responses, *b* 10th year responses, *c* Years until 95% of long run effect reached

We can see that four mass points do not seriously affect the results. Moreover, to add normal distributed random effects in wages does only increase wage elasticities at the intensive margin slightly. The normal distribution assumption on random effects in the last specification increases all elasticities slightly. To conclude, the results seem relatively robust to exactly what type of random effects are included, although we have seen that it is important not to ignore the effect of unobserved heterogeneity altogether.

[22] The normal distribution has standard deviation according to the residuals in the wage regression. For each individual the likelihood value is averaged over the 30 draws.

[23] A more sophisticated method for normal distributed unobserved heterogeneity would be to estimate not only the standard deviation but also the correlation between the three parameters. However, as this slows down the optimization process considerably, we have chosen a more restricted estimation where α_{0i} and β_{0i} are perfectly correlated (but differ with respect to standard errors), whereas μ_i is assumed perfectly uncorrelated with the two other random effects. We have used an adjusted version of a program provided by Haan and Uhlendorff (2006) in which Halton draws, as suggested by Train (2009), are used.

4.6 Conclusion and Discussion

This is the first application to extend the static job choice model to an intertemporal setting by allowing for unobserved heterogeneity and true state dependence. The job choice model offers an interesting framework for an attempt to distinguish between state dependence in preferences, capturing the effect that last periods labor supply choice might affect future evaluation of the consumption/leisure choice, and state dependence in the opportunity measure capturing how last periods leisure choice affect job opportunities next period.

We have developed and estimated the intertemporal model for the subgroup of women in couples. The same framework can be used for men and single women. However, women in couples differ from the other groups by a wider variation in working hours. As our data source is administrative and not based on survey information, the measure of working hours is relatively crude, and only contains for instance contractual working hours and registered overtime work. As the vast majority of Norwegian males work at least full time, we do not think it is much to gain to estimate the model jointly for the couple, although possible. In order to predict the full labor supply responses to a tax reform, one would have to consider the complete set of wage earners. Still, we believe that the results for women in couples can provide an interesting example of the dynamics of labor supply responses to a policy intervention.

We have focused on explaining sources for unobserved heterogeneity and state dependence within a highly structural framework. The advantage of more structure is that we have a framework to explain and distinguish the sources of persistence and adjustments. In a less structural and more ad hoc framework one will usually attain a better fit, but it is questionable to which degree one has captured the deeper underlying preference parameters. Moreover, with weaker assumptions regarding the functional form of unobserved heterogeneity, state dependence and serial correlation it is hard if not impossible to identify the different effects. The assumptions regarding time independent unobserved heterogeneity and a special restrictive form of serial correlation are, however, certainly strong and could influence the estimated effect of state dependence (see e.g. Prowse, 2012). We have therefore conducted several alternative specification to show how the model fit and the wage elasticities change.

The model can be characterized as quite flexible when it comes to the five possible choices of labor supply, coupled with a detailed incorporation of the tax and benefit system. Remember that our model is myopic since it assumes that individuals optimize on a period to period basis. The dynamics follow from that this periods decision will influence next periods evaluation of preferences and choice set, although this is not explicitly taken into account by the individual. A life cycle model is a more sophisticated framework to approach, but it is not clear to which degree individuals indeed are long-sighted rational, and it can be questioned how appropriate this type of model is to analyze relatively short sighted tax reforms. Moreover, as one might gain more reality in terms of a life cycle perspective, one looses the possibility for other details necessary to defend a model which can be used for policy makers in practice.

Appendix

Table 4.A1: Pooled Heckman 2-Stage wage regression: Women in couples

	ln(Wage)		Participation	
Experience	0.0164***	(0.0003)	0.0023	(0.0040)
Experience squared	-0.0003***	(0.0000)	-0.0008***	(0.0001)
Low education	-0.0968***	(0.0022)	-0.2082***	(0.0201)
High education	0.2791***	(0.0016)	0.4424***	(0.0207)
Non-western origin	-0.1192***	(0.0038)	-0.9994***	(0.0234)
Residence in metropolitan area	0.0733***	(0.0013)	-0.0781***	(0.0142)
Year dummies (base: 2003)				
Year 2004	0.0459***	(0.0017)	0.0606**	(0.0189)
Year 2005	0.0820***	(0.0017)	0.1878***	(0.0193)
Year 2006	0.1243***	(0.0017)	0.2488***	(0.0196)
Year 2007	0.1768***	(0.0017)	0.3279***	(0.0201)
Education category (base: "unknown")				
General	0.0206***	(0.0052)	0.6999***	(0.0323)
Human, Art	-0.0624***	(0.0055)	0.4348***	(0.0396)
Education	-0.0736***	(0.0054)	0.9016***	(0.0423)
Social, Law	0.0557***	(0.0062)	0.8511***	(0.0614)
Business	0.0284***	(0.0052)	0.8295***	(0.0329)
Technology	0.0677***	(0.0054)	0.8294***	(0.0387)
Health	-0.0809***	(0.0052)	1.0256***	(0.0331)
Primary	-0.0049	(0.0095)	0.3533***	(0.0787)
Service	-0.0267***	(0.0064)	0.6344***	(0.0499)
Constant	4.6834***	(0.0068)	1.5323***	(0.0653)
Exclusion restrictions				
No. children under age 3			-0.1002***	(0.0206)
No. children under age 6			-0.0879***	(0.0174)
No. children under age 12			-0.3349***	(0.0098)
Wealth (in NOK 10,000)			-0.0000**	(0.0000)
Husband's net income (in NOK 10,000)			-0.0024***	(0.0001)
Mills lambda	0.0273***	(0.0058)		
Number of observations	110,001			
Number of individuals	26,631			

Standard errors in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Chapter 5

Concluding Remarks

This dissertation is a contribution to the empirical literature on labor supply responses to income taxation along two main avenues. First, I have demonstrated how simulations from a structural labor supply model (ex-ante) can be compared to results from panel data approaches with identification of earnings responses from a tax reform period (ex-post). Second, I have, in both methodological frameworks, analyzed the dynamics of labor supply responses over time, in order to investigate the time frame of people's responses.

In Chapter 2, the particular tax changes associated with the Norwegian tax reform in 2006 was utilized in order to compare simulation predictions from a structural labor supply model with results from a panel data evaluation measuring income responses. My main finding is that simulations from the structural labor supply model yield net-of-tax elasticity estimates that are close to the elasticity estimated on the basis of the panel data; ranging from 0.05 to 0.09 for the structural model and from 0.04 to 0.055 for the panel data analysis.

In Chapter 3 and 4, the panel data approach and the structural labor supply model were both extended in order to analyze the dynamics of labor supply responses. Interestingly, results for both methods suggest that long run responses in labor supply are about twice as large as short run (first year) responses. Moreover, the strong autoregressive income component in the panel data approach suggests that state dependence in working hours might indeed be an important explanation for observed inertia in labor supply and earnings.

In Chapter 3, I find, in accordance to the previous chapter, small responses in earnings with elasticities of 0.06/0.12 in the short/long run in the dynamic panel data model compared to 0.06 for the static panel data. In order to present the policy implications of these estimates, I simulated a change in different hypothetical surtax schedules. I find that up to 40 percent of the surtax cuts (marginal tax rate reduction from 0.5 to 0.45) can be financed through additional generated income.

In Chapter 4, I incorporated state dependence in preferences and job opportunities in a structural labor supply model for women in couples. I estimated wage elasticities of 0.13-0.28 in the short and long run for a permanent shock in wages. I further presented a set of hypothetical policy changes to demonstrate the predicted slower responses in labor supply. It can be noted that in accordance to Haan (2010), I find that the long run wage elasticities predicted from the intertemporal model are similar to the wage elasticities of the static model presented in Chapter 2.

In this dissertation, I have concentrated on, what has become, two workhorses within empirical public economics: The discrete choice labor supply model and the panel data method of estimating the elasticity of taxable income.

The structural discrete choice model is well-known among practitioners of labor supply analysis. Once the model is estimated, it can be used to simulate ex-ante changes in the tax schedule or in the wage rate. Discrete choice models, as opposed to the traditional continuous labor supply models, can easily deal with complex tax-benefit systems with non-linear and non-convex budget sets, and is therefore a tool for giving suggestions to policy makers in practice. The static structural labor supply model reflects individual heterogeneity in the population and is estimated on cross-sectional data. In this dissertation I have focused on a job choice model developed in Statistics Norway, which offers a more coherent theoretical foundation for the discrete choice models.

Still, the discrete choice model is founded on a highly theoretical framework, which clearly simplifies individuals behavior at the labor market. This in addition to strict assumptions on the error terms. Although the intertemporal model in Chapter 4 looses some of the restrictions with concern to the time dimension of responses, there are clearly underlying assumptions regarding unobserved heterogeneity and the structure of state dependence. It can be difficult to test the identifying assumptions, and one can accordingly be critical to the simulation results the model provide. Quasi-experimental studies offer on the other hand a more straightforward identification strategy.

The literature on elasticity of taxable income (ETI) has attained considerable attention the last decades. The ETI literature was motivated by the wish to capture the full set of responses to income taxation, by directly measuring responses in taxable income. After Feldstein (1995) pioneered the ETI panel data method of grouping individuals in treatment and control groups based on their pre-reform income level, the approach has been extended by Auten and Carroll

(1999) and Gruber and Saez (2002). Although there have been some attempts and criticism of the conventional approach, the method can still be considered as “state of the art”.

In this dissertation I utilized the ETI methodology in order to analyze responses in labor income for wage earners. Although the focus on wage earners is not broad enough to capture the full responses of income taxation, it is of prior interest to decompose the responses, also from a methodological point of view. Moreover, the focus on wage earners is particular interesting in the present analysis in order to compare with predicted responses from the labor supply model. One of the main arguments in Chapter 2 was that a simple comparison of average wage elasticities from the labor supply model with average net-of-tax rates from the ETI approach is misleading. ETI estimates are derived from specific tax reforms, and therefore measure the average effects for the individuals treated by the reform. The non-linearity of the labor supply model, on the other hand, implies different responses along the income distribution.

Clearly, tax reforms are seldom designed as random experiments, in which individuals are selected to the old or to the new tax schedule by chance. A typical tax reform involves changes in the tax rate for a particular range of income. In the panel framework I have adopted here, one utilizes the individuals’ income position in a base year period in order to construct an instrument which group individuals in treated and less or not treated by the reform. From the results, it seems evident that this instrument is strong. However, there are, not surprisingly, many possible pitfalls when trying to estimate the effect of a tax reform based on comparing outcomes for individuals in the lower and upper part of the income distribution. The challenge is to estimate a realistic reduced form model which is in accordance with both tax reform periods, as well as periods where no reforms occurred. As the tax rate is only one of many aspects influencing individual income development over time, there exists a risk of misinterpretation. This, in particular, if characteristics unobserved to the researcher differ between the people affected and not affected by the reform.

In order to alleviate these problems, I have estimated the models over periods both with and without tax reforms. Moreover, utilizing a data source with the unique combination of tax return data and detailed individual characteristics puts me in an advantageous position. As the ETI methodology has been shown to be highly dependent on sample restrictions (see e.g. Giertz, 2010), my robustness tests for both the static and dynamic model suggest a much more robust result. One reason for this might be that I focus on a more homogenous group of individuals, namely wage earners at the mid-level and high-end of the income distribution. This group is associated with less income mobility, which suggests that the tax rate instrument is more appropriate. It is moreover promising that the conventional ETI model provides similar estimates to the dynamic panel data model in the short run.

Another aspect with the ETI approach is that it is not obvious through which channels the tax responses work. I have focused on explanations such as real responses in labor supply through individual choice of working hours and effort. One might, however, think of other channels. There is for instance a literature which investigates how unions prevail on wage moderation as a reaction to tax progressivity (see e.g. Lockwood et al., 2000). Moreover, Piketty et al. (2011) suggest that individual bargaining responses might be important.

Accordingly, for the dynamic panel data analysis, it is not clear why there is a strong autoregressive component in income which leads to slower tax responses. It could be due to state dependence in labor supply (analogously to the structural model in Chapter 4), but one could also think of other sources mentioned in Chapter 3 such as cost of changing working hours/job, information lag and habit persistence. As it can be hard to distinguish these effects in a reduced form model, one prospect for further research would be to incorporate more of these possible channels in a structural framework.

In general, one should be aware of that both the structural discrete choice model and the ETI framework are supply side models in which the demand side of the labor market is toned down. Possible extensions of the structural model could be to include search costs and to elaborate further on the job opportunity measure as a demand side restriction. Further, it would be interesting to align the structural labor supply model with the underlying theoretical framework of the ETI models. The job choice model is a suitable starting point for extension in terms of a broader set of responses in which one for instance chooses jobs which differ with respect to earnings.

Although I have focused on the time frame in an intertemporal structural model and in a dynamic panel data model, both methods can still be placed within the concept of static type of models. This is in contrast to dynamic models with a life cycle perspective. A life cycle model is a more sophisticated framework to approach, but it is not clear to which degree individuals indeed are long-sighted rational, and it can be questioned how appropriate this type of model is in analyzing relatively short sighted tax reforms. Moreover, as one might gain more reality in terms of a life cycle perspective, one looses the possibility for other details necessary to defend a model which can be used for policy makers in practice. However, it would clearly be advantageous to incorporate aspects of the life cycle models such as forward looking expectations in the "static" models I have assessed in this dissertation.

I have concentrated on comparing the structural labor supply model with outcome from the

ETI literature. Another more straightforward comparison would be to analyze working hours directly over a reform period. I have not chosen to follow this course, because of lack of detailed panel data on working hours, and also because I wanted to draw on the experience from estimating panel data in the literature on elasticity of taxable income. Clearly, similar challenges would apply when analyzing working hours in order to set up a reduced form model which is appropriate both in tax-reform and non-tax-reform periods. An analysis along these lines would, however, certainly be interesting both for validating the structural model in a more straightforward fashion, and for learning more about the anatomy of the estimated elasticity of taxable income elasticities.

It would further be useful to validate the structural choice model for the lower income levels and at the extensive margin. More research is also needed on how responses differ along the income scale, and to which extent the ETI estimates can be used in a simulation context of prospective policy changes.

Bibliography

Aaberge, R., J. K. Dagsvik, and S. Strøm (1995). Labor Supply Responses and Welfare Effects of Tax Reforms. *Scandinavian Journal of Economics 97*(4), 635–659.

Aarbu, K. O. and T. O. Thoresen (2001). Income Responses to Tax Changes- Evidence from the Norwegian Tax Reform. *National Tax Journal 54*(3), 319–335.

Aasnes, J., J. Dagsvik, and T. Thoresen (2007). The Norwegian Tax-Benefit Model System LOTTE. In A. Gupta and A. Harding (Eds.), *Modelling our Future: Population Ageing, Health and Aged Care, International Symposia in Economic Theory and Econometrics*, pp. 513–518. Elsevier Science.

Anderson, T. W. and C. Hsiao (1981). Estimation of Dynamic Models with Error Components. *Journal of the American Statistical Association 76*(375), 598–606.

Angrist, J. D. and J.-S. Pischke (2010). The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics. *The Journal of Economic Perspectives 24*(2), 3–30.

Auten, G. and R. Carroll (1999). The Effect of Income Taxes on Household Income. *Review of Economics and Statistics 81*(4), 681–693.

Bækgaard, H. (2010). Earned Income Response to Tax Changes in Denmark. *Ministry of Finance Working Paper Series*.

Bingley, P. and I. Walker (1997). The Labour Supply, Unemployment and Participation of Lone Mothers in In-Work Transfer Programmes. *The Economic Journal 107*(444), 1375–1390.

Bloemen, H. G. and A. Kapteyn (2008). The Estimation of Utility-Consistent Labor Supply Models by Means of Simulated Scores. *Journal of Applied Econometrics 23*(4), 395–422.

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), 878–889.

Blundell, R. (2006). Earned Income Tax Credit Policies: Impact and Optimality: The Adam Smith Lecture, 2005. *Labour Economics 13*(4), 423–443.

Blundell, R. (2010). Comments on: Michael P. Keane 'Structural vs. Atheoretical Approaches to Econometrics'. *Journal of Econometrics 156*(1), 25–26.

Blundell, R., A. Duncan, J. McCrae, and C. Meghir (2000). The Labour Market Impact of the Working Families' Tax Credit. *Fiscal Studies 21*(1), 75–104.

Blundell, R., M. Francesconi, and W. van der Klaauw (2011). Anatomy of Policy Reform Evaluation: Announcement and Implementation Effects. *IZA Discussion Paper No. 6050*.

Blundell, R. and T. MaCurdy (1999). Labor Supply: A Review of Alternative Approaches. In O. C. Aschenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3, pp. 1559–1695. Elsevier.

Blundell, R. and A. Shephard (2012). Employment, Hours of Work and the Optimal Taxation of Low-Income Families. *The Review of Economic Studies 79*(2), 481–510.

Bond, S. (2002). Dynamic Panel Data Models: A Guide to Microdata Methods and Practice. Technical report, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.

Brewer, M., A. Duncan, A. Shephard, and M. J. Suarez (2006). Did Working Families' Tax Credit Work? The Impact of In-Work Support on Labour Supply in Great Britain. *Labour Economics 13*(6), 699–720.

Burtless, G. and J. A. Hausman (1978). The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Tax Experiments. *Journal of Political Economy 86*(6), 1103–30.

Cai, L. (2010). Work Choices of Married Women: Drivers of Change. *Visiting Researcher Paper, Productivity Commission, Canberra*.

Cai, L., G. Kalb, Y.-P. Tseng, and H. Vu (2008). The Effect of Financial Incentives on Labour Supply: Evidence for Lone Parents from Microsimulation and Quasi-Experimental Evaluation. *Fiscal studies 29*(2), 285–325.

Carroll, R. and W. Hrung (2005). What Does the Taxable Income Elasticity say about Dynamic Responses to Tax Changes? *American Economic Review 95*(2), 426–431.

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

Chetty, R. (2009b). Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods. *Annual Review of Economics 1*(1), 451–488.

Chetty, R. (2012). Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply. *Econometrica 80*(3), 969–1018.

Chetty, R., J. N. Friedman, T. Olsen, and L. Pistaferri (2011). Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records. *The Quarterly Journal of Economics 126*(2), 749–804.

Chetty, R., A. Guren, D. Manoli, and A. Weber (2011). Are Micro and Macro Labor Supply Elasticities Consistent? A Review of Evidence on the Intensive and Extensive Margins. *American Economic Review 101*(3), 471–75.

Cogan, J. F. (1981). Fixed Costs and Labor Supply. *Econometrica 49*, 945–963.

Creedy, J. and G. Kalb (2005). Discrete Hours Labour Supply Modelling: Specification, Estimation and Simulation. *Journal of Economic Surveys 19*(5), 697–734.

Creedy, J., G. Kalb, and R. Scutella (2006). Income Distribution in Discrete Hours Behavioural Microsimulation Models: An Illustration. *Journal of Economic Inequality 4*(1), 57–76.

Dagsvik, J., Z. Jia, T. Kornstad, and T. Thoresen (2013). Theoretical and Practical Arguments for Modeling Labor Supply as a Choice Among Latent Jobs. *forthcoming in Journal of Economic Surveys*.

Dagsvik, J. K. (1994). Discrete and Continuous Choice, Max-Stable Processes, and Independence from Irrelevant Attributes. *Econometrica 62*(5), 1179–1205.

Dagsvik, J. K. and Z. Jia (2012). Labor Supply as a Discrete Choice Among Latent Jobs. *Discussion Paper No. 709, Statistics Norway*.

Dagsvik, J. K. and A. Karlström (2005). Compensating Variation and Hicksian Choice Probabilities in Random Utility Models that are Nonlinear in Income. *Review of Economic Studies 72*(1), 57–76.

Dagsvik, J. K. and S. Strøm (2006). Sectoral Labour Supply, Choice Restrictions and Functional Form. *Journal of Applied Econometrics 21*(6), 803–826.

Deaton, A. (2010). Instruments, Randomization, and Learning about Development. *Journal of Economic Literature 48*(2), 424–55.

Eckstein, Z. and K. I. Wolpin (1989). Dynamic Labour Force Participation of Married Women and Endogenous Work Experience. *The Review of Economic Studies 56*(3), 375–390.

Eissa, N. and H. W. Hoynes (2006). Behavioral Responses to Taxes: Lessons from the EITC, and Labor Supply. *Tax Policy and the Economy 20*, 73–110.

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.

Feldstein, M. (1999). Tax Avoidance and the Deadweight Loss of the Income Tax. *Review of Economics and Statistics 81*(4), 674–680.

Giertz, S. H. (2010). The Elasticity of Taxable Income During the 1990s: New Estimates and Sensitivity Analyses. *Southern Economic Journal 77*(2), 406–433.

Gottfried, P. and D. Witczak (2009). The Responses of Taxable Income Induced by Tax Cuts-Empirical Evidence from the German Taxpayer Panel. *IAW Discussion Paper 57, Institut für Angewandte Wirtschaftsforschung*.

Gruber, J. and E. Saez (2002). The Elasticity of Taxable Income: Evidence and Implications. *Journal of Public Economics 84*(1), 1–32.

Haan, P. (2006). Much Ado About Nothing: Conditional Logit vs. Random Coefficient Models for Estimating Labour Supply Elasticities. *Applied Economics Letters 13*(4), 251–256.

Haan, P. (2010). A Multi-State Model of State Dependence in Labor Supply: Intertemporal Labor Supply Effects of a Shift from Joint to Individual Taxation. *Labour Economics 17*(2), 323–335.

Haan, P., D. Kemptner, and A. Uhlendorff (2012). Bayesian Procedures as a Numerical Tool for the Estimation of Dynamic Discrete Choice Models. *IZA Discussion Paper No. 6544*.

Haan, P. and V. Steiner (2005). Distributional Effects of the German Tax Reform 2000- A Behavioral Microsimulation Analysis. *Schmollers Jahrbuch: Journal of Applied Social Science Studies 125*(1), 39–49.

Haan, P. and A. Uhlendorff (2006). Estimation of Multinomial Logit Models with Unobserved Heterogeneity using Maximum Simulated Likelihood. *Stata Journal 6*(2), 229–245.

Haan, P. and A. Uhlendorff (2012). Intertemporal Labor Supply and Involuntary Unemployment. *Empirical Economics*, 1–23.

Hansen, J. and X. Liu (2011). Estimating Labor Supply Responses and Welfare Participation: Using a Natural Experiment to Validate a Structural Labor Supply Model. *IZA Discussion Paper No. 5718*.

Hansson, Å. (2007). Taxpayers' Responsiveness to Tax Rate Changes and Implications for the Cost of Taxation in Sweden. *International Tax and Public Finance 14*(5), 563–582.

Hausman, J. A. (1985). The Econometrics of Nonlinear Budget Sets. *Econometrica 53*(6), 1255–82.

Heckman, J. (1981). Statistical models for discrete panel data models. In C. F. Manski and D. McFadden (Eds.), *Structural Analysis of Discrete Data with Econometric Applications*, pp. 114–175. MIT Press Cambridge, MA.

Heckman, J. and B. Singer (1984). A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data. *Econometrica 52*(2), 271–320.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica 47*(1), 153–162.

Heckman, J. J. (1993). What Has Been Learned About Labor Supply in the Past Twenty Years? *The American Economic Review 83*(2), 116–121.

Heckman, J. J. (2010). Building Bridges Between Structural and Program Evaluation Approaches to Evaluating Policy. *Journal of Economic Literature 48*(2), 356–398.

Heckman, J. J. and S. Urzua (2010). Comparing IV with Structural Models: What Simple IV can and cannot Identify. *Journal of Econometrics 156*(1), 27–37.

Holmlund, B. and M. Söderström (2011). Estimating Dynamic Income Responses to Tax Reform. *The BE Journal of Economic Analysis & Policy 11*(1).

Hyslop, D. R. (1999). State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labor Force Participation of Married Women. *Econometrica 67*(6), 1255–1294.

Imbens, G. W. (2010). Better LATE than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009). *Journal of Economic Literature 48*(2), 399–423.

Imbens, G. W. and J. M. Wooldridge (2009). Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature 47*(1), 5–86.

Keane, M. and R. Rogerson (2012). Micro and Macro Labor Supply Elasticities: A Reassessment of Conventional Wisdom. *Journal of Economic Literature 50*(2), 464–76.

Keane, M. P. (2010a). A Structural Perspective on the Experimentalist School. *The Journal of Economic Perspectives 24*(2), 47–58.

Keane, M. P. (2010b). Structural vs. Atheoretic Approaches to Econometrics. *Journal of Econometrics 156*(1), 3–20.

Keane, M. P. (2011). Labor Supply and Taxes: A Survey. *Journal of Economic Literature 49*(4), 961–1075.

Keane, M. P. and K. I. Wolpin (2007). Exploring the Usefulness of a Nonrandom Holdout Sample for Model Validation: Welfare Effects on Female Behavior. *International Economic Review 48*(4), 1351–1378.

Kiss, Á. and P. Mosberger (2011). The Elasticity of Taxable Income of High Earners: Evidence from Hungary. *MNB Working Papers No. 11, The Central Bank of Hungary*.

Kleven, H. J. and E. A. Schultz (2011). Estimating Taxable Income Responses using Danish Tax Reforms. *EPRU Working Paper, University of Copenhagen*.

Labeaga, J. M., X. Oliver, and A. Spadaro (2008). Discrete Choice Models of Labour Supply, Behavioural Microsimulation and the Spanish Tax Reforms. *Journal of Economic Inequality 6*(3), 247–273.

LaLonde, R. J. (1986). Evaluating the Econometric Evaluations of Training Programs with Experimental Data. *The American Economic Review 76*(4), 604–620.

Lindsey, L. B. (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.

Lockwood, B., T. Sløk, and T. Tranæs (2000). Progressive Taxation and Wage Setting: Some Evidence for Denmark. *The Scandinavian Journal of Economics 102*(4), 707–723.

Massarrat-Mashhadi, N. and C. Werdt (2012). Estimating Dynamic Income Responses to Tax Changes: Evidence from Germany. *Discussion Paper No. 22, Freie Universität Berlin*.

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behaviour. In P. Zarembka (Ed.), *Frontiers in Econometrics*, pp. 105–142. Academic Press New York.

McFadden, D. and K. Train (2000). Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics 15*(5), 447–470.

McFadden, D. L. (1984). Econometric Analysis of Qualitative Response Models. In Z. Griliches and M. D. Intriligator (Eds.), *Handbook of Econometrics*, Volume 2, pp. 1395–1457. Elsevier.

Meghir, C. and D. Phillips (2008). Labour Supply and Taxes. *Mirrlees Review, Institute for Fiscal Studies*.

Moffitt, R. (1984). The Estimation of a Joint Wage-Hours Labor Supply Model. *Journal of Labor Economics 2*(4), 550–66.

Moffitt, R. A. 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*, pp. 193–234. New York: Russel Sage Foundation.

OECD (2011). *Growing Income Inequality in OECD-Countries. What Drives it and How Can Policy Tackle it*. Report.

Okamura, K. and N. Islam (2009). State Dependence in Labour Force Participation of Married Women in Japan. *IRISS Working Paper Series No. 17*.

Piketty, T., E. Saez, and S. Stantcheva (2011). Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities. *NBER Working Paper No. 17616*.

Pronzato, C. D. (2012). Comparing Quasi-Experimental Designs and Structural Models for Policy Evaluation: The Case of a Reform of Lone Parental Welfare. *IZA Discussion Paper No. 6803*.

Prowse, V. (2012). Modeling Employment Dynamics with State Dependence and Unobserved Heterogeneity. *Journal of Business & Economic Statistics 30*(3), 411–431.

Saez, E., J. Slemrod, and S. H. Giertz (2012). The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review. *Journal of Economic Literature 50*(1), 3–50.

Skrondal, A. and S. Rabe-Hesketh (2004). *Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models*. Chapman & Hall/CRC.

Slemrod, J. (1992). Do Taxes Matter? Lessons from the 1980's. *American Economic Review 82*(2), 250–256.

Slemrod, J. (1995). Income Creation or Income Shifting? Behavioral Responses to the Tax Reform Act of 1986. *The American Economic Review 85*(2), 175–180.

Slemrod, J. and W. Kopczuk (2002). The Optimal Elasticity of Taxable Income. *Journal of Public Economics 84*(1), 91–112.

Soest, A. v. (1995). Structural Models of Family Labor Supply: A Discrete Choice Approach. *Journal of Human Resources 30*(1), 63–88.

Statistics Norway (2003). *Labour Force Survey 2001*. Official Statistics of Norway (NOS D 748).

Statistics Norway (2005). *Income Statistics for Persons and Families 2002-2003*. Official Statistics of Norway (NOS D 338).

Statistics Norway (2006). *Wage Statistics 2005*. Official Statistics of Norway (NOS D 362).

Statistics Norway (2009). *Quality Report on the Norwegian Structure of the Earnings Survey 2006*. Report No. 2009/20.

Statistics Norway (2013). Statbank Table 04903. `https://www.ssb.no/en/utdanning/statistikker/barnehager`.

Statistics Norway (2013). Statbank Table 08517. `https://www.ssb.no/en/arbeid-og-lonn/statistikker/aku`.

Thoresen, T. O., J. Aasness, and Z. Jia (2010). The Short-Term Ratio of Self-Financing of Tax Cuts: An Estimate for Norway's 2006 Tax Reform. *National Tax Journal 63*(1), 93–120.

Thoresen, T. O. and A. Alstadsæter (2010). Shifts in Organizational Form under a Dual Income Tax System. *FinanzArchiv: Public Finance Analysis 66*(4), 384–418.

Thoresen, T. O., E. E. Bø, E. Fjærli, and E. Halvorsen (2012). A Suggestion for Evaluating the Redistributional Effects of Tax Changes: With an Application to the 2006 Norwegian Tax Reform. *Public Finance Review 40*(3), 303–338.

Todd, P. E. and K. I. Wolpin (2006). Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility. *The American Economic Review 96*(5), 1384–1417.

Train, K. E. (2009). *Discrete Choice Methods with Simulation* (Second ed.). Cambridge University Press.

Van Soest, A., M. Das, and X. Gong (2002). A Structural Labour Supply Model with Flexible Preferences. *Journal of Econometrics 107*(1), 345–374.

Weber, C. (2011). Obtaining a Consistent Parameter of the Elasticity of Taxable Income using Difference-in-Differences. *Working Paper, University of Michigan*.

Wooldridge, J. M. (2005). Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity. *Journal of Applied Econometrics 20*(1), 39–54.

List of Tables

List of Figures

English Summary

Labor supply responses to tax changes are a core issue in public economics, reflected by numerous estimates from different methodological approaches. Relationships between labor supply and taxes in a microeconomic and microeconometric perspective are often discussed based on two categories of research. The first line of research refers to estimation of structural labor supply models for simulation purposes (ex-ante). A second main method is based on analyses of observations before and after a realized policy reform (ex-post).

Along both empirical methods (as well as in the theoretical literature) it is typically assumed that individuals react instantly to changes in the tax schedule, either at the time of implementation of a policy reform or already at the time when the policy reform is announced. Possible reasons for slower responses in labor supply, and consequently the need for a model framework which incorporates the time dimension, include information lag, habit persistence and costs associated with changes in hours worked (Holmlund and Söderström, 2011). Moreover, when analyzing the broader set of labor supply responses through earnings, some responses such as increased effort, responsibility and human capital investment might lead to a rise in earnings only after some time has passed. State dependence, defined as the causal effect of experiencing an event on preferences, prices or constraints relevant to future choices (Heckman, 1981), represents another source of observed inertia in people's reactions to policy changes.

This dissertation contributes to the literature along two main avenues. First, I demonstrate how simulations from a static discrete choice model can be validated by panel data methods estimating the elasticity of taxable earnings (Chapter 2). Second, I analyze the dynamics of labor supply responses to tax changes over time both with regards to panel data approaches (Chapter 3) and incorporated in a structural labor supply framework (Chapter 4).

I focus on responses in working hours and labor income for wage earners, and use Norwegian register and survey data provided by Statistics Norway.

In Chapter 2, I demonstrate how the standard elasticity of taxable income (ETI) methodology can be used to validate predictions from a structural labor supply model, by analyzing the

Norwegian tax reform in 2006. The structural labor supply model discussed, following e.g. Dagsvik and Jia (2012), and Dagsvik et al. (2013), is related to the discrete model of Soest (1995). Traditional methods in the ETI literature are used (following e.g. Gruber and Saez, 2002), utilizing the panel structure of data to obtain individual measures of income growth, and employing instrumental variable techniques to obtain measures of change in the net-of-tax rate. To facilitate comparison, I use the discrete choice model to simulate the effects of the 2006 tax reform on hours of work, and use predicted income levels to obtain a comparable estimate for income elasticity with respect to the net-of-tax rate. Both methods suggest small responses (net-of-tax elasticities below 0.1) in labor supply and earnings to the particular tax reform of consideration.

In Chapter 3, I analyze the dynamics of earnings responses to tax changes by exploiting substantial exogenous variation in the two-tier surtax schedule for labor income over a period of 14 years (1995-2008). I adopt the dynamic panel data framework by Holmlund and Söderström (2011), and compare with estimates from the conventional static panel approach by Gruber and Saez (2002). The estimated magnitude of the elasticities of earnings with respect to net-of-tax are modest, about 0.06 for the three-year conventional static panels, and 0.06/0.12 in the short/long run for the preferred dynamic specification. I find that the long run responses to tax rate changes in labor supply are about twice as large as the short run responses due to a strong autoregressive effect in earnings. Supplementary conducted simulations suggest that up to 40 percent of the revenue loss due to tax rate cuts in the surtax schedule can be self-financed in the long run by inducing additional generated income.

In Chapter 4, I examine the effect of state dependence to policy change responses over time within the framework of an intertemporal structural discrete choice model. I extend the static discrete choice microsimulation model described in Chapter 2 to an intertemporal setting. One advantage of the applied theoretical framework is that it allows me to distinguish between state dependence in preferences and state dependence in the set of job opportunities. I estimate the model for the subset of female wage earners in couples and find significant positive state dependence both at the intensive and extensive margin. This causes sluggish labor supply responses to various changes in the wage rate and the tax schedule. Within the first year of a permanent policy change the simulation results suggest that about 50 percent of the full labor supply response can be expected, whereas 90 percent of the response is reached after about 4 years both at the intensive and at the extensive margin.

German Summary

Der Einfluss der Einkommensteuer auf das Arbeitsangebot ist ein zentrales Thema in der Finanzwissenschaft, welches sich in zahlreichen Schätzungen verschiedener methodischer Ansätze widerspiegelt. Der Zusammenhang zwischen Arbeitsangebot und Steuern wird aus mikroökonomischer und mikroökonometrischer Perspektive oft in zwei Forschungskategorien betrachtet. Ein Forschungsansatz bezieht sich auf die Schätzung struktureller Arbeitsangebotsmodelle für Simulationszwecke (ex-ante). Eine zweite Methode basiert auf Beobachtungsanalysen vor und nach einer Reform (ex-post). Die Elastizität des zu versteuernden Einkommens (ETI) ist dabei ein zentrales Konzept, bei dem die Auswirkungen von Grenzsteuersätzen auf Einkommen betrachtet werden.

Innerhalb beider empirischer Ansätze wird davon ausgegangen, dass Personen unmittelbar auf Änderungen im Steuertarif reagieren; entweder bei der Umsetzung einer Reform oder bereits zum Zeitpunkt der Ankündigung der Reform. Informationsverzögerungen, Verharren im Gewohnten und der Aufwand zur Änderung der Arbeitsstunden sind mögliche Ursachen für verzögerte Reaktionen auf das Arbeitsangebot. Daraus folgend bedarf es eines Modellkonzeptes, welches die zeitliche Dimension berücksichtigt (Holmlund and Söderström, 2011). Die Zustandsabhängigkeit, die als kausaler Zusammenhang zwischen bisherigem Arbeitsangebot und zukünftigen Präferenzen, Löhnen und Budgetbeschränkungen definiert wird (Heckman, 1981), stellt einen weiteren Trägheitsfaktor bei der Reaktion auf Reformen dar.

Diese Dissertation trägt zur vorhandenen Literatur anhand zweier Vorgehensweisen bei. Im ersten Schritt zeige ich, wie Simulationen eines statischen, diskreten Wahlmodells anhand von Methoden für Paneldaten validiert werden können (Kapitel 2). Im zweiten Schritt analysiere ich die Dynamik des Arbeitsangebots bei steuerlichen Änderungen über Paneldatenansätze (Kapitel 3), sowie über ein strukturelles Arbeitsangebotsmodell (Kapitel 4). Ich konzentriere mich auf Auswirkungen auf die Arbeitszeit und Arbeitseinkommen von Lohnempfängern, bei Verwendung norwegischer Registrierungs- und Befragungsdaten.

In Kapitel 2 zeige ich, wie die Standard-ETI-Methodik verwendet wird, um Mikrosimulationen eines strukturellen Arbeitsangebotsmodells anhand der norwegischen Steuerreform von 2006 zu validieren. Das verwendete strukturelle Arbeitsangebotsmodell (u.a. Dagsvik and Jia, 2012, Dagsvik et al., 2013) ist angelehnt an das diskrete Modell von Soest (1995). Anhand traditioneller Methoden der ETI-Literatur (nach e.g. Gruber and Saez, 2002) werden Instrumentvariablen eingesetzt, um Änderungen von Grenzsteuersätzen zu messen. Um einen Vergleich zu ermöglichen, verwende ich das diskrete Wahlmodell, um die Auswirkungen der Steuerreform auf die Arbeitsstunden zu simulieren. Weiterhin nutze ich die simulierten Einkommen, um eine Schätzung der Einkommenselastizität in Bezug auf den Grenzsteuersatz zu erhalten. Beide Methoden deuten auf geringe Effekte (Elastizitäten unter 0,1) auf das Arbeitsangebot und die Einkommen bei der betrachteten Steuerreform hin.

In Kapitel 3 analysiere ich die Dynamik der Einkommenseffekte auf Steueränderungen. Hierfür werden wesentliche Variationen im zweistufigen Zusatzsteuertarif für Arbeitseinkommen über eine Periode von 14 Jahren (1995-2008) verwendet. Ich wende das dynamische Paneldatenkonzept von Holmlund and Söderström (2011) an und vergleiche die Ergebnisse mit denen des konventionellen, statistischen Panelansatzes nach Gruber and Saez (2002). Die geschätzten Einkommenselastizitäten in Bezug auf den Netto-Grenzsteuersatz ($1-$Grenzsteuersatz) sind gering - etwa 0,06 für die 3-jährigen statischen Panels, und 0,06 / 0,12 in kurz- bzw. langfristiger Zeitspanne für die dynamische Spezifikation. Ich stelle fest, dass die langfristigen Arbeitsangebotseffekte, aufgrund starker autoregressiver Einflüsse auf das Einkommen, etwa doppelt so groß wie die kurzfristigen Reaktionen sind. Zusätzlich durchgeführte Simulationen weisen darauf hin, dass sich langfristig bis zu 40 Prozent der Einkommenseinbüßen bei Steuerkürzungen durch zusätzlich generiertes Einkommen selbstfinanzieren lassen.

In Kapitel 4 untersuche ich den Einfluss der Zustandsabhängigkeit auf das zeitliche Reaktionsverhalten bei fiskalpolitischen Änderungen. Hierzu erweitere ich das statische Mikrosimulationsmodell aus Kapitel 2 zu einem intertemporalen Arbeitsangebotsmodell. Ich berechne das Modell mit Daten weiblicher Lohnempfänger die in einer Partnerschaft leben. Die Ergebnisse deuten auf eine signifikante, positive Zustandsabhängigkeit hin, sowohl in Bezug auf Präferenzen als auch auf Arbeitsmöglichkeiten. Auf Änderungen im Lohn- und Steuertarif reagiert das Arbeitsangebot demzufolge nur zögernd. Die Simulationsergebnisse weisen darauf hin, dass ungefähr 50 Prozent der gesamten Arbeitsangebotseffekte innerhalb des ersten Jahres nach einer Stundenlohnänderung zu erwarten sind, während 90 Prozent der Effekte nach ca. 4 Jahren erreicht werden.