Income Responses to Tax Changes. Reconciling Results of Quasi-Experimental Evaluation and Structural Labor Supply Model Simulation

by

Thor O. Thoresen* and Trine E. Vattø†

Preliminary draft, May 10, 2012

Abstract
Labor responses to tax changes are often discussed by employing a structural labor supply model to simulate responses on working hours. An alternative source of information on behavioral response is comparison of labor incomes before and after changes in the tax schedule (as a tax reform), employing a quasi-experimental identification strategy. This paper brings these two strands of the literature together by using them to discuss income responses of reductions in marginal tax rates at high income levels, which means that results of the two approaches can be compared and interpreted in relation to each other. Both sources of information suggest that the responses of the 2006 tax reform are rather modest.

Keywords: Labor supply, Behavioral effects, Tax Responses, Discrete choice structural model, Elasticity of taxable income
JEL classification: H21, H24, H31, J22

*Statistics Norway, Oslo, Norway and CESifo
Corresponding author**:** Thor O. Thoresen, Research Department, Statistics Norway, P.O.B.8131 Dep., N-0033 Oslo, Norway. Phone: +47-21094819, fax: +47-21090040, email: thor.olav.thoresen@ssb.no
† Statistics Norway, Oslo, Norway and PhD student at the Free University of Berlin, Germany, email: tev@ssb.no

1. Introduction

Individual labor supply and income responses to tax changes is 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, by simulation of tax changes applying a static structural labor supply model and by response estimates obtained from analysis of panel data, comparing incomes before and after a particular tax change. The standard procedure under the first line of research is to estimate a static structural labor supply model. From observations of households and individuals' consumption and connections to the labor market, typically working hours, one can either fit a labor supply function directly or one can estimate a utility function, see reviews of the literature in Blundell and MaCurdy (1999) and Keane (2011). The parameter estimates can in turn be used to simulate effects of changes in the tax system.

The second main method to obtain information about relationships between income and taxes often centres the attention on income responses, which represents broader behavioral responses (than effects on working hours or labor market participation). Identification of response estimates typically apply the difference-in-differences estimator or related econometric techniques, measuring treatment effects by utilizing that tax reforms can be seen as defining quasi-experiments in the sense that they generate net-of-tax rate changes along the income scale, often producing substantial tax changes for some tax-payers, whereas others are more or less unaffected. A key concept is the elasticity of taxable income (ETI), which measures the response in taxable income for a change in the net-of-tax rate. Saez, Slemrod and Giertz (2009; 2012) survey the literature.

Even though there are some examples of studies which discuss experimental evidence in relation to results from structural models, see LaLonde (1986), Eissa and Hoynes (2004), Todd and Wolpin (2006) and Blundell (2006), we have seen less "cross-bearings" of results from the ETI-studies[1] and structural labor supply model simulations. Blundell (2006) argues that "simple difference in difference evaluations can be valuable for validating the specification of more fragile microeconometric models" (p. 425). But how can results from a structural labor supply model be compared to estimates derived from the quasi-experimental method in meaningful way? This study brings these two strands of the literature together by using both methodological approaches to discuss how responsive tax-payers are to a particular change in tax rates. By doing this we offer a practical suggestion to facilitate comparison of results across methods. Further, as the response of income from tax changes is a measure which holds a key position in the public policy debate, cross-bearing of results of the two empirical approaches is essential in the search for valid measures and for the

[1] By ETI-studies we refer to reduced form studies developed the last couple of decades (after initial contributions by Lindsey (1987) and Feldstein (1995)), focusing on income responses and using "experimental" empirical identification strategies.

understanding of what such measures express. Obviously, it is reassuring if both sources of information point to similar response magnitudes, but given that the two approaches pick up different effects, response estimates will not be identical as there are remaining sources to disparate outcomes. These reasons for differences are spelled out in the present paper.

We focus on the response of wage earners at the high end of the income distribution, which follows from the identification of the estimates of the quasi-experimental approach, exploiting the reductions in top marginal tax rates of the 2006 tax reform to derive earned income elasticities. Traditional methods of 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. These results are compared to results from a structural labor supply model simulation, facilitated by estimation of a discrete choice model. To facilitate comparison with the ETI-results, instead of only reporting wage elasticities, we simulate the effects on hours of work of the specific tax reform, and use predicted income levels to obtain an estimate for income elasticity with respect to the net-of-tax, which is the key measure of the ETI-literature.

The paper is organized as follows. In Section 2 we present the two methodological approaches to obtain tax response estimates, followed by presentations of results in Section 3. In Section 4 we bring the results together and discuss what they convey about Norwegian tax responsiveness. Section 5 concludes the paper.

2. Empirical models for income and tax relationships

One will find a whole range of different response estimates 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,[2] tax simulation based on the structural discrete choice labor supply model and estimation of the elasticity of taxable income under a quasi-experimental identification strategy. Given that estimation of structural labor supply models often involve severe econometric challenges,[3] see reviews in Blundell and MaCurdy (1999) and Keane (2011), reduced form estimation based on the difference-in-differences estimator may represent a rather straightforward empirical technique for the practitioner of public finance. However, besides that identification methods rely on rather strong assumption, see e.g., Moffitt and Wilhelm (2000), a main limitation of the ETI-approach is that the "treatment effect" must be interpreted in terms of the specific tax change under consideration, and therefore is less informative about effects of other policy changes. But even though there are empirical concerns regarding both sources of information on tax responses,

[2] Chetty et al. (2011) refer to this type of evidence as steady-state elasticities.
[3] It can be argued that the discrete choice version of structural modelling represents is more practical than the conventional continuous approach, based on marginal calculus. The structural labor supply model associated with Hausman becomes very complicated in the case when more general and flexible model specifications are used, see Bloemen and Kapteyn (2008).

they provide an opportunity for cross-bearing of empirical results, which is illustrated by the present analysis.

Recently we have witnessed discussions in the literature concerning interpretations and advantages of "structural modeling" versus "reduced form" approaches, see for instance Chetty (2009), Deaton (2010), Imbens (2010), Keane (2010) and Heckman (2010). As emphasized by Chetty (2009), the ETI approach cannot easily be placed according to the two stereotype classifications, since these elasticities share important characteristics with both strands of the literature.[4] For instance, similar to structural models the ETI framework departs from an underlying utility maximizing behavior and renders precise statements about welfare implications. The identification strategy shares, however, important similarities with reduced form or experimental studies. In this section we present the main characteristics of the two methods to derive response estimates. First we present a discrete choice labor supply model and then next we describe how tax response estimates can be derived from panel data analysis.

2.1 Choice of working hours based on a discrete choice model formulation

Discrete choice models of labor supply based on the random utility modeling approach have gained widespread popularity, mainly because it is much more practical than the conventional continuous approach based on marginal calculus; see Van Soest (1995) for an outline of standard discrete choice model. 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.

The labor supply model applied here is based on a version of the discrete choice model formulation, where the agents are assumed to make choice with respect to "jobs"; see Dagsvik and Strøm (2006), Dagsvik and Jia (2012), and Dagsvik et al. (2012). Each job is characterized by a discrete set of hours (as in the traditional model), but several jobs might be characterized with the same working hours. In addition to consumption and leisure, the individual is assumed to have preferences over jobs which are unobserved for the researcher. This means that the utility function of the household can be seen as $U(C,h,z)$, where *z* = 1, 2,…, refer to market opportunities (jobs) and $z=0$ refers to the nonmarket alternative. The utility function is assumed to have multiplicative structure, $U(C,h,z)=v(C,h)+\varepsilon(z)$. where *v*(·) is a positive deterministic function and the random unobserved components $\varepsilon(z)$ are dependent on job z in addition to unobserved individual characteristics. We assume that the random components are i.i.d. extreme value distributed with c.d.f. $\exp\left(\exp\left(-x\right)\right)$ for positive *x*. The distribution assumption implies independence of irrelevant alternatives (IIA), which is a common assumption in the discrete choice labor supply literature.

Let $\psi(h) = v\left(f\left(hw, I\right), h\right)$ be the representative utility of jobs with hours of work *h*, a given wage rate *w* and non-labor income *I*. In a more general set-up, one may allow wages to vary across jobs, see Dagsvik and Jia (2012), but here we will let the wage depend on individual characteristics, only.[5] 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 $B(h)$. There is only one nonmarket 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 shall choose job *z* can be expressed as

$$\text{(2.1)} \quad P\left(v(f(hw,I),h)+\varepsilon(z)=\max_{x\in D\cup\{0\}}\max_{k\in B(x)}(v(f(xw,I),h)+\varepsilon(k))\right)$$

$$=\frac{\exp\psi(h)}{\sum_{x\in D,}\sum_{z\in B(x)}\exp\psi(x)+\exp\psi(0)}.$$

However, $\psi(h) = v\left(f\left(hw, I\right)h\right)$ is defined as the representative utility of a job with working hours *h*. In order to derive an expression for the probability for choosing any job within $B(h)$, we sum over all the alternatives within $B(h)$, that is,

$$\text{(2.2)} \quad \varphi(h)=\sum_{z\in B(h)}\frac{\exp(\psi(h))}{\sum_{x\in D,}\sum_{z\in B(x)}\exp(\psi(x))+\exp(\psi(0))}=\frac{\exp(\psi(h))m(h)}{\exp(\psi(0))+\sum_{x\in D}\exp(\psi(x))m(x)},$$

When $h = 0$ we get

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

[5] The simplification we shall follow is that the agent considers an individual specific wage rate, thus with no variation across jobs. Instead we address the mean offered wage rate, also introducing a random effect to account for unobserved heterogeneity in wage rate opportunities. Introducing random effects in the wage equation may also be seen as loosen the somewhat restrictive form of the conditional logit model, referred to as the IIA restriction (Dagsvik et al., 2012).

(2.3) $$\varphi(0)=\frac{\exp(\psi(0))}{\exp(\psi(0))+\sum_{x\in D}\exp(\psi(x))m(x)},$$

Let θ define the total number of jobs available to the individual. Then one can define $g(h)$ as the fraction of jobs available to the agent with offered hours of work equal to h, $g(h)=m(h)/\theta$. We shall call $\theta g(h)$ the opportunity measure and $g(h)$ the opportunity distribution. When inserting the opportunity measure into the expressions for probabilities, we obtain

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

and

(2.5) $$\varphi(0)=\frac{\exp\left(\psi(0)\right)}{\exp\left(\psi(0)\right)+\theta\sum_{x\in D,x>0}\exp\psi(x)g(x)}.$$

The resulting expression is a choice model that is analogous to a multinomial logit model with representative utility terms $\{\psi(h,w)\}$, weighted by the frequencies of available jobs, $\{m(h)=\theta g(h)\}$. Unfortunately, $m(h)$ is not observable, but under the assumption that $g(h)$ is uniformly distributed over individuals with peaks at part and full time work, and by assuming that theta is individual specific and depending on the individuals education, the model can be estimated.

Appendix A shows how $v(C,h)$ and the wage rate is specified, and present the estimation results for single males, single females, and separately for males and females in couples (married/cohabiting), which are utilized in the simulation of behavioral responses to the tax changes, presented in Section 3.

2.2 Utilizing direct observations of income growth

The approach followed in much of the ETI-literature departs from an underlying utility maximizing behavior similar to what is seen in the standard labor supply literature above (Feldstein,1999; Saez, Slemrod and Giertz, 2012). Individuals are assumed to maximize a utility function which increases in consumption(C) and decreases in taxable income (q), subject to a budget constraint described by $C=\left(1-\tau\right)q+R$, where τ is the marginal tax rate (at a linear segment of the tax schedule), and R is virtual income. In the present context we define *q* to be earned taxable income, defined as wage rate (*w*) times working hours (*h*). Thus, this formulation suggests closer relationship to the part of the structural labor supply literature which 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).[6]

Whereas standard labor supply approaches usually focus on the choice of *h* given an individual-specific wage rate, a main advantage of the ETI-approach is that it opens up for a broader range of responses to changes in marginal tax rates captured by the income response, as denoted by Feldstein (1995). In this study we focus on the real responses in wage income capturing possible responses in hours and wages. This can be identified as we use changes in the tax schedule for labor income, and as we look at responses in non-deductible taxable labor income.

We adopt the measure of the elasticity of income with respect to changes in the net-of-tax rate, defined by $e = \frac{1-\tau}{q}\frac{\delta q}{\delta(1-\tau)}$. Panel data covering a period of net-of-tax rate variation across individual and across time (often covering a tax reform) has been the main data source for identification of ETI-estimates. If 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, $\log(1-\tau_{it})$, an individual effect, μ_i, and an error term, ξ_{it},

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

the basic framework for identification in the ETI literature is various estimations of a first differenced version of (1), using panel data for two periods and eliminating the individual effect, μ_i,

(2.7) $$\Delta \log q_i = \kappa + \lambda \Delta \log(1-\tau_i) + \xi_i.$$

The reliability of results rests upon carefully framed empirical designs for identification of the key parameter, including controls for effects from observed and unobserved characteristics. A main methodological identification challenge (of λ) has been the endogeneity of the tax variable, which has led to estimation of (2.7) by IV techniques, for instance employing the difference in differences estimator, grouping the individuals into treated and non-treated based on pre-reform income levels. Feldstein (1995) is an example of this.[7] Many post-Feldstein studies employ a closely related instrument, using (for the net-of-tax rate variable) the change in net-of-tax rates according to first period income as the excluded variable in the IV estimation, see Auten and Carroll (1999) and Gruber

[6] This structural model specification thus deviates from the standard discrete choice model (Van Soest, 1995) and the discrete choice model presented above, in which estimation is carried out directly on the utility specification.

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

and Saez (2002). Thus, this line of research relies heavily on methods commonly used in the "experimentalist" or "program evaluation" literature.[8] As tax reforms often involve reductions or increases in maximum marginal tax rates, and small or no changes at lower income levels,[9] the treatment and controls groups follow from their income level. Thus, we are far from the randomized trial interpretation of results that many studies seek to obtain.

The ETI literature focuses on effects that are similar to the average treatment effect of the treated. In other words, if we let a parameter δ be a zero-one indication of being treated (experiencing net-of-tax rates changes or not), as in Feldstein (1995), one identifies $E(\lambda|\delta_{it}=1)$. According to Blundell and MaCurdy (1999), this parameter is subject to conventional sample selection biases and cannot be used to simulate policy responses. In so far as we think this is too pessimistic, as we suggest that ETI estimates can be used for validation of predictions from structural models (as also noted by Blundell, 2006), such measures are valuable from a tax policy perspective as they contain crucial information about behavioral effects and efficiency effects of tax changes (Feldstein, 1999; Chetty, 2009).

Even though this type of panel data analysis is characterized as non-structural according to standard typologies, the specification of the reduced form is helped by important lessons from the structural labor supply literature. For instance, a carefully designed empirical approach would need to address income effects. Similar to Blomquist and Selin (2009) we construct virtual incomes by procedures similar to the approaches seen in the labor supply literature, based on piece-wise linear approximations to the budget constraint (see Burtless and Hausman, 1978). Virtual income will be expressed by the difference between paying the marginal tax on overall labor income, $\tau_{it}q_{it}$, and the actual taxes paid, given by $\upsilon(q_{it})$. This difference will be positive for a progressive tax system with tax allowances. In addition, since q_{it} only captures labor income, we will include non-labor income I_{it} as exogenously given.

$$(2.8) \qquad R_{it} = I_{it} + \left(\tau_{it}q_{it} - \upsilon_{it}(q_{it})\right)$$

In non-labor income we will include untaxed transfers, such as the child benefit and other social transfers in addition to net of tax capital income. For couples, non-labor income includes the income of the spouse.

Appendix B provides a more detailed description of how this type of model can be estimated, given the data we have had access to.

[8] There are conceptual challenges when categorizing different studies. Two tags that are used to define non-structural studies are "program evaluation" (Imbens and Wooldridge, 2009) and "experimentalist" (Keane, 2010).

[9] At least this has been the case both in 1992 and 2006 in Norway.

3. Tax response estimates

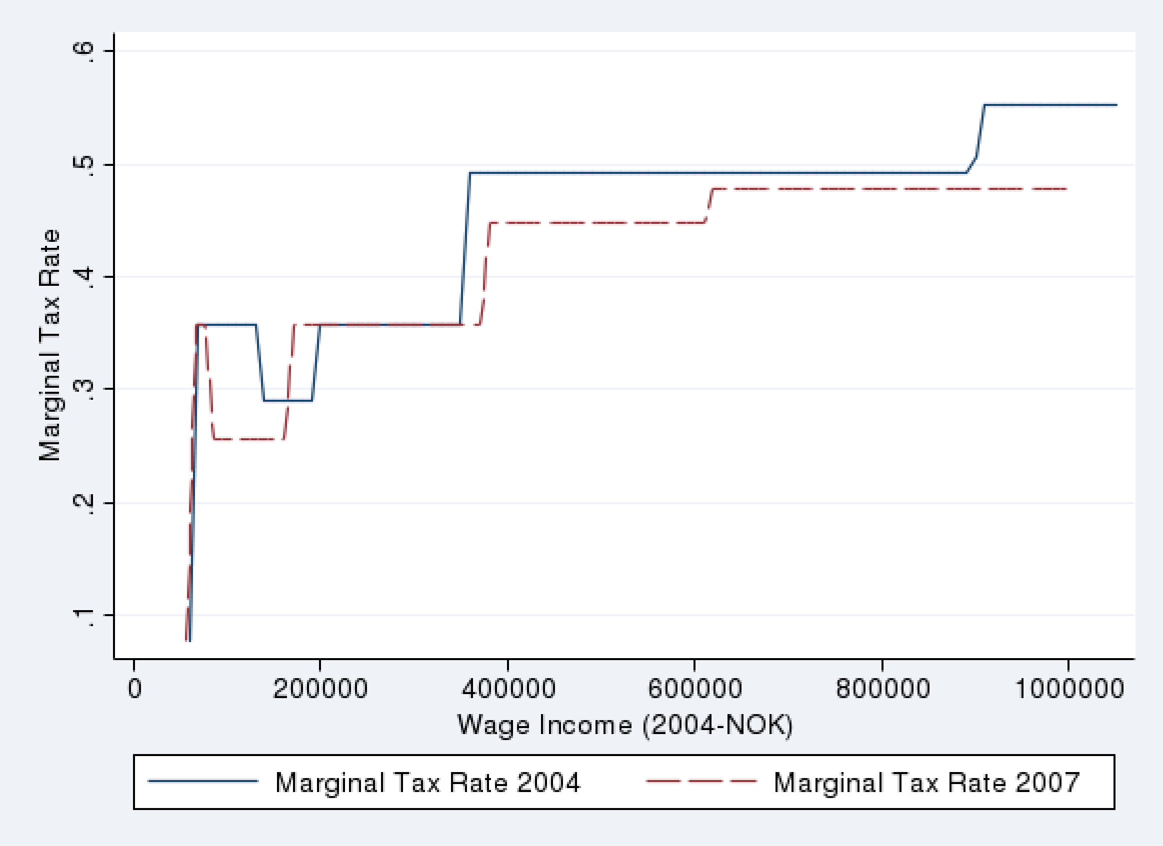
In this section we probe deeper into the “cross-bearing” of the results of the two methodologies, to discuss the empirical content of the two sources of information, and ultimately, assess to what extent they provide similar estimates of tax-payers income responses to tax changes. The change in marginal tax rates on wage income of the Norwegian tax reform of 2006 is used to illustrate the effects. After providing some institutional background on the tax reform, we present the evidence of the panel data quasi-experimental approach, and then next, these results are contrasted to the predictions of the labor supply model.

3.1 The reductions in marginal tax rates by the tax reform of 2006

Norway has a “dual income tax” system, enacted in a 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 throughout the Nordic countries in the early 1990s. The Norwegian version had a flat 28 percent tax rate levied on corporate income, capital 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, introducing a shareholder income tax, and most importantly in the present context, by cutting labor income marginal tax rates.

The tax reform was gradually implemented in the years 2005 and 2006. Figure 1 reflects 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 at NOK354,300 at a rate of 13.5 percent, and the second tier of 19.5 percent applied to income in excess of NOK906,900. The reform implied that the maximum marginal tax rate fell from 55.3 to 47.8 percent, but became effective at a lower level.

Figure 1. Reductions in marginal tax rates according to the tax reform of 2006



3.2 Evidence from panel data estimations

We closely follow the conventional approach in the ETI-literature, see e.g. Gruber and Saez (2002), where changes in net-of-tax rates are instrumented by the tax change for a constant individual income level. More details on the empirical specification and sample restrictions are presented in Appendix B. The exogenous variation in this study is the Norwegian tax reform, which (as already noted) was gradually implemented during 2005–2006. As the tax instrument is based on the initial period income and the dependent variable is growth in income, a control for mean reversion and drifts in the income distribution is necessary. Auten and Caroll(1999) included therefore the initial income as an additional explanatory variable, and Gruber and Saez (2002) extended this approach by allowing for a piecewise linear function of initial income. We adopt this approach by including 10 linear splines or a three degree polynomial of initial income.

The main data source is the Income Statistics for Persons and Families (Statistics Norway, 2006a), a register-based data set which cover the complete Norwegian population, with data from income tax returns as a main component. The panel dimension can be easily exploited as each individual is coded with a personal identification number. We restrict the data set to wage earners in the age group 25–62, defined as having labor income as their main income source and exclude students and individuals with positive self-employed income, pensions or unemployment benefits.

We use six overlapping 3-year panels over the period 2000–2008. The reason why we have chosen to include a wider dataset, outside the reform period 2005–2006, is to improve the estimates for the control variables, in particular the mean reversion control. All wage earners with income in the upper 2/3 of the income distribution (equals about NOK250,000 in 2004) in the base year (the first year in the respectively 3-year panel period) are included in our main analysis. There are two reasons for excluding the lower income levels. Firstly, we are mainly interested in the effect of decreased surtax rates, which affect only about 1/3 of the wage earners. Secondly, the mean reversion problem is especially severe for individuals with initially low income, which makes this group less appropriate as a control group.

Table 1 reports the results of the 2SLS regressions. In the first two columns, 10 splines of log income are included, whereas in the third and fourth column a third degree centered polynomial of log income is included as mean reversion control. Specification (2) and (4) include a control for virtual and non-labor income. Although results (in general) are sensitive to the inclusion of the mean reversion control, there is only a minor difference between the estimates including 10 splines or a third degree polynomial of base year income. The elasticity of labor income with respect to net-of-tax is estimated to about 0.05–0.06 without income effect and 0.03–0.04 after the income effect is controlled for. The estimated virtual and non-labor income elasticity is small and negative, as expected.

The estimated net-of-tax elasticities are very small when compared to most other ETI studies, the literature more often reports estimates in the range 0.2–0.8. According to Saez, Slemrod and Giertz (2012), estimates from the U.S. (after Feldstein, 1995) range from 0.12 to 0.40. One reason for the low response reported here may is that we measure real responses in labor income for the restricted group of wage earners, which means that we do not capture any altered deduction behavior and assumingly less short-sighted tax planning. Moreover, our estimates might be less influenced by drifts in the income distribution (unrelated to the tax reform) as the wage distribution was relatively stable (or followed a linear trend) over the period of consideration. Our ETI-results will therefore, if correctly specified, cover the changes in hours of work in addition to changes in effort (changes in hourly wage). The time span is probably too short to capture more general effects on education attainment etc. Still, the responses suggest that Norwegian wage earners are less responsive to the tax changes imposed by the 2006 reform.

Table 1. 2SLS Regression results for all wage earners

	Mean reversion control			
	10 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)
Non-labor income elasticity		-0.0091***		-0.0105***
		(0.0012)		(0.0012)
Splines of (income/median)	Yes	Yes		
3 degree polynomial of (income/median)			Yes	Yes
Number of observations	4,933,291	4,331,276	4,933,291	4,331,276

Note: All regressions include control variables for gender, wealth, age, age squared, married, number of children under and above 6, newborn, residence in Oslo/ dense populated area, non-west origin, years of education, dummies for education area, income shifting control and year dummies. Full regression output is reported in table B1, in Appendix B.

We have also divided the sample into four groups, single females, single males, females in couple and males in couple, to have a closer look at responses for specific groups and to facilitate closer comparison with the simulation results of the structural model estimation. A third degree polynomial is used as a mean reversion control and we exclude the income control in order to compare with the results from the structural model. The results of Table 2 suggest that the responses are similar in the four groups of wage earners, ranging between 0.030 for single males and 0.045 for males in couple. For females, the elasticity of earned income is estimated to 0.034 for singles and 0.041 for women in couple. Note that although the estimates are small, they are all highly statistically significant, due to a large number of observations.

Table 2. 2SLS Regression results for groups of wage earners

	3 degree polynomial, no income elasticity			
	Single females	Single males	Females, couple	Males, couple
Net-of-tax elasticity	0.0340***	0.0298***	0.0410***	0.0448***
	(0.0061)	(0.0059)	(0.0049)	(0.0031)
Number of observations	576,232	959,151	1,109,651	2,287,960

Note: All regressions include control variables for wealth, age, age squared, married, number of children under and above 6, newborn, residence in Oslo/ dense populated area, non-west origin, years of education, dummies for education area, income shifting control and year dummies. Full regression output is reported in table B2.

3.2. Results of simulations based on labor supply model

Next, we show how we can derive estimates of comparable net-of-tax elasticities from a labor supply model simulation. The discrete choice structural model is estimated by using information on hours work from the Labor Force Survey (Statistics Norway, 2005) and income data from the Income Statistics for Persons and Families (Statistics Norway, 2006a) for 2004 (a pre-reform year). Four separate models, for men in couple, women in couple, single women and single men, are estimated. In

Appendix A results of the labor supply model estimations are presented, including the results of the estimations of wage rate equations.[10]

Given that the model deviates from the standard discrete choice models in terms of accounting for differences in number of job options across individuals and peaks in the distribution of working hours, θ and $g(h)$, respectively, see Equation (2.4), it is worth noting that the number of job options is increasing in education and that the full-time peak is reflected by a parameter estimate well above 1.

Before addressing results of simulations of the income elasticity with respect to the net-of-tax rate, we present standard wage elasticities of the estimated model. The uncompensated wage elasticities are estimated by increasing gross hourly wage by one percent, and simulate the percentage change in predicted hours worked for each individual. The average elasticity for each group is given in Table 3. 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, and show modest elasticities, in the range 0.07–0.27.

Table 3. Gross wage elasticity estimates

	Gross Wage Elasticities		
	Total	Extensive Margin	Intensive Margin
Males in couple	0.16 (0.xxxx)	0.003 (0.xxxx)	0.16 (0.xxxx)
Single males	0.08 (0.xxxx)	0.008 (0.xxxx)	0.07 (0.xxxx)
Females in couple	0.36 (0.xxxx)	0.087 (0.xxxx)	0.27 (0.xxxx)
Single females	0.25 (0.xxxx)	0.065 (0.xxxx)	0.18 (0.xxxx)

Note that wage elasticities are calculated by using a percentage increase in gross wage, and are not directly comparable to the net-of-tax elasticities from the ETI-literature. But a simple example can describe their similarities. Let $W = w(1-\overline{\tau})$ where W is net wage, w is gross wage and $\overline{\tau}$ is the average tax level. Now assume that we have a simple two-step tax system where the two tax rates are τ_1 and τ_2 respectively. Imagine that basic allowances are absent and let the cut-offs for each tax bracket be kept constant. Now, the net-of-tax rate is increased by 1 percent for each threshold. Under the assumption that we do not have any basic allowances, if $(1-\overline{\tau})$ is increased by 1 percent (for a contant w), this is identical to the gross wage, *w*, being increased by one percent (as long as the average tax rate is kept constant). Moreover, note that as *w* is considered to be constant at the individual level in the structural model, a percentage change in hours is identical to a percentage change in labor income. So, under these simplifications, we have

[10] As further elaborated upon in Appendix A regressions account for selectivity bias for females, not for males. The individual wage is represented by the predicted wage rate, with an additional random effect. In practice, the random effect is accounted for by making 30 draws, from the measure of the error term variance, and subsequently applying maximum simulated likelihood by computing expected values for the individual log likelihood function across the 30 draws.

$$(2.11)\quad \frac{\partial h/h}{\partial W/W}=\frac{\Delta h/h}{\Delta(1-\tau)/(1-\tau)\mid w}=\frac{\Delta wh/wh}{\Delta(1-\tau)/(1-\tau)}.$$

However, when increasing gross wage by one percent, both the average and the marginal tax rate may increase, such that the net wage increase could be less than 1 percent in magnitude. The most important complication is, however, that the structural model is nonlinear, and since there are no identifiable quasi-experiment where all wage earners face the same net-of-tax change, the two measures will never be immediately comparable. A first step to obtain comparable measures of net-of-tax rate elasticities from the labor supply model is to simulate the effects of the 2006-reform on working hours.[11] These results are shown in Table 4 for the four groups of wage earners.[12] As for the wage elasticities, predicted hours under the pre- and post-reform schedules are based on the estimated probability distribution for each individual.

Table 4. Predicted hours pre- and post-reform

	Pre-reform	Post-reform	Difference
Males in couple	38.90 (0.xxx)	39.26 (0.xxx)	0.91 %
Single males	38.94 (0.xxx)	39.13 (0.xxx)	0.48 %
Females in couple	36.04 (0.xxx)	36.19 (0.xxx)	0.45 %
Single females	37.09 (0.xxx)	37.21 (0.xxx)	0.34 %

Next, to obtain an overall estimate of the ETI for the structural model, we simply regress predicted growth in labor income on the change in net-of-tax rate, as in the ETI-literature. Growth in labor income is identical to growth in predicted hours for an individual specific wage rate, and the net-of-tax rate is instrumented by similar methods as in the ETI-literature, using the change in net-of-tax for constant (predicted) initial labor income (predicted pre reform hours times the individual's constant wage rate) as the instrument[13]. The estimated elasticties are reported in Table 5.

Table 5. Estimated net-of-tax elasticities for the structural model

Growth in hours/labor income	Net-of-tax rate elasticity	Std. Error
Males in couple	0.103	(0.xxxx)
Single males	0.059	(0.xxxx)
Females in couple	0.052	(0.xxxx)
Single females	0.048	(0.xxxx)

[11] The 2007 brackets are deflated to a 2004 income level by using the median wage growth over the period.
[12] Random draws are used to determine the specific predicted hour choice pre and post reform.
[13] As in the experimental approach, the regression is restricted to individuals with predicted pre reform income in percentile 33 or above. Note that the predicted income distribution is very similar to actual income distribution due to the inclusion of random wage residuals.

We see that the comparable net-of-tax rate elasticities are somewhat lower than the wage elasticties, in the range between 0.05–0.10. Moreover, although the estimated wage elasticities are clearly higher for women (0.18-0.27 for females versus 0.07-0.16 for males), the net-of-tax elasticity results suggest that females are about equally or less responsive than males. This follows from the model's predictions of stronger similarity between female and male wage elasticities at the high end of the income distribution.[14]

4. Reconciling the evidence

This study brings together two approaches which are widely used and accessible for practical policy evaluation. It is, however, important to keep in mind that the two types of models are based on different assumptions and frameworks. Let us therefore first review some of the main differences, such as discrete/continuous choice, responses through working hours/total labor income, the underlying time frame and more generally, the distinction between a structural approach with simulation and an experimental approach.

Firstly, the structural model we have estimated is based on discrete choice instead of marginal optimization. In the discrete choice structural models we estimate a certain probability distribution for different options of working hours.[15] There are different practical alternatives which can be employed in the simulation of such models, but in the present model the simulation of alternative policies implies that the overall probability distribution is altered as the economic conditions change. This means that an individual who choose a part-time job in a pre-reform year will also be affected by a tax reform where only the surtax rates are altered. In the ETI-literature it is instead (somewhat simplified) assumed that individuals are either treated or not treated by the reform, and typically an individual working part-time will be seen as a non-treated individual in this context. The ETI-literature is based on marginal optimization and therefore is more similar to the continuous hours structural labor supply models (the line of research often associated with the Hausman model).

Secondly, the models differ in the type of responses that tax changes induce. As already emphasized, the ETI literature may cover a whole range of responses, including tax planning and tax avoidance, as it typically focus on total taxable income. In our study we approach a more narrow focus on wage earners' responses in labor income (hourly wage times hours). Still, we should capture responses in both working hours and wage rate. It has been argued that the assumption of a fixed exogenously given individual wage in the structural labor supply model is too strict. In the ETI-

[14] Wage elasticities for each decile of the wage distribution uncover how the model predicts responses to vary over the income distribution.

[15] Recall that the model is a "job choice" model, which is turned into a choice between different categories of hours of work.

literature it has been argued that also the wage rate can be seen partly as a choice variable for the individual as he or she may alter the wage through increased efforts per hour or job shifts.

Thirdly, the methods probably differ somewhat to the time frame. The structural model is a static model where a new long run steady-state immediately is attained. In the experimental method, on the other hand, we use the ad hoc choice of 3-year spans. As well as the structural model might be inappropriate for describing short-term responses, it is not obvious how such results can be compared to the time framing of the experimental method.

Lastly, in the structural approach one simulates the responses of a tax change based on a cross-sectional model with a highly theoretical framework, whereas in the experimental approach one estimates elasticities based on direct observation of income before and after the tax change. 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 as we seek to estimate the deep underlying structural parameters. However, as the model may be too simple or suffer from misspecification, it may be tempting to argue that the experimental approach is a test of how well the structural model performs. In this view the experimental approach would "uncover" the true responses. However, this is not necessary straightforward. Ideally in an experimental approach, we would namely 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, a main practical problem in the conventional experimental 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, unrelated to the tax reforms. One may therefore raise serious concerns to what extent one is able to reveal unbiased estimates of the ETI.[16]

Despite the major differences in the methodological framework for the two models, the estimates are reasonable similar. In Table 6 we restate the comparable results of the structural model and the experimental panel data estimation.

Table 6. Comparison of net-of-tax rate elasticity estimates from structural labor supply model simulations and direct observations of income (ETI framework)

	Structural Model	Panel Data
Males in couple	0.103 (0.xxxx)	0.045 (0.0061)
Single males	0.059 (0.xxxx)	0.030 (0.0059)
Females in couple	0.052 (0.xxxx)	0.041 (0.0049)
Single females	0.048 (0.xxxx)	0.034 (0.0031)

The net-of-tax elasticities are small in both the structural and the quasi-experimental model, in the range 0.03-0.1. It is somewhat surprising that the structural model actually predicts somewhat larger

[16] In another paper (Dagsvik, Thoresen and Vattø, 2012) we discuss a method for estimating the ETI which use alternative estimation techniques.

responses than the experimental approach, although it only covers the responses in working hours. It might, however, be important to notice that the structural model is estimated on actual in contrast to contractual hours of work (see Appendix A). This means that we allow for responses which does not necessarily correspond to a job shift in the sense that you shift from a contractual full time job to a contractual over time job, but it could mean that you “shift” to a contractual full time job in which you take on more responsibility (still at the same working place) where you know you often need to work overtime (possibly unpaid), but with a corresponding rise in monthly pay. This means that the structural model capture a broader set of responses than would could be expected from responses in contractual hours.

It is often acknowledged in the labor supply literature that high income individuals are typically less responsive to working hours, as there is a natural or institutional limit to working hours per week. In the ETI framework on the other hand one typically finds large elasticities also for high income individuals, which can be explained by other margin of responses such as income shifting and tax planning behaviour, and through effort decisions and thereby productivity per hour. Our estimates are much smaller than typically found in the literature, possibly because we focus on the real responses in labor income, in contrast to taxable income. Also, we have chosen to 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, capital earners and individuals with less strong attachment to the labor market.

For both methods we estimate the uncompensated elasticities. It is uncommon and complicated to report compensated elasticities in the discrete structural labor supply literature, see however Dagsvik and Karlstrøm (2005) for a method to derive measures of compensated effects in discret choice models. The income effect is typically estimated to be small in the ETI-literature such that it is often assumed that the compensated and uncompensated elasticities are similar (see e.g. Saez, Slemrod and Giertz, 2009).

In general, it might be argued that the Norwegian institutional setting produces smaller elasticities; the argumentation presented by Slemrod and Kopczuk (2002) may be used in support for Norwegians being less responsive.

5. Summary

Empirical estimates of labor supply responses to changes in taxation can be derived from various methodological approaches. Two main sources of information are simulations of responses based on estimated structural labor supply models and quasi-experimental panel data estimations, comparing incomes before and after a particular tax change. The former approach typically report elasticities of predicted hours worked with respect to gross hourly wage rate. The latter approach uses tax reforms for defining quasi-experiments, in the sense that they generate net-of-tax rate changes along the income scale, often producing substantial tax changes for some tax-payers, whereas others are more or less unaffected. The key concept is the elasticity of taxable income with respect to net-of-tax rate (1-marginal tax rate). In this paper we have shown that wage elasticities from the structural model are typically not directly comparable to the net-of-tax elasticities. Instead the respective tax reform is simulated and regressed on instrumented net-of-tax rates in order to obtain comparable results.

Our main finding is that both sources of information give rather low response estimates. As a "cross-bearing" of the information from these two sources of information about responses, the evidence presented here point in the same direction. Norwegian median and high income wage earners react rather modestly to tax changes.

References

Aarbu, K.O. and T.O. Thoresen (2001), Income Responses to Tax Changes – Evidence from the Norwegian Tax Reform. *National Tax Journal*, 54, 319–35.

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

Blomquist, S. and H. Selin (2010), Hourly Wage Rate and Taxable Labor Income Responsiveness to Changes in Marginal Tax Rates. *Journal of Public Economics*, 94, 878–889.

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, 395–422.

Blundell, R. (2006), Earned Income Tax Credit Policies: Impact and Optimality. The Adam Smith Lecture 2005. *Labour Economics,* 13, 423–44.

Blundell, R. and T. MaCurdy (1999), "Labor Supply: A Review of Alternative Approaches". In O.C. Ashenfelter and D. Card (eds.): *Handbook of Labor Economics*, Vol. 3A, Amsterdam: North-Holland, 1559–1695.

Burtless, G. and J. Hausman (1978), The Effect of Taxes on Labor Supply. *Journal of Political Economy*, 86, 1103–1130.

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

Chetty, R., A. Guren, D. Manoli and A. Weber (2011), Are Micro and Macro Labor Supply Elasticities Consisten? A Review of Evidence on the Intensive and Extensive Margins. *American Economic Review*, 101, 471–475.

Deaton, A. (2010), Instruments, Randomization, and Learning about Development. *Journal of Economic Literature,* 48, 424–455.

Dagsvik, J. and A. Karlström (2005), Compensating variation and Hicksian choice probabilities in random utility models that are nonlinear in income. *Review of Economic Studies*, 72, 57–76.

Dagsvik, J.K., and Z. Jia (2012): Labor Supply as a Job Assignment Problem, Manuscript, Statistics Norway.

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

Dagsvik, J.K., T.O. Thoresen and T.E. Vattø (2012), Measures to Estimate the Elasticity of Taxable Income Revisited. Paper in progress, Statistics Norway.

Dagsvik, J.K., Z. Jia, T. Kornstad and T.O. Thoresen (2012), Theoretical and Practical Arguments for Modeling Labor Supply as a Choice among Latent Jobs. CESifo Working Papers No. 3708, January 2012, Center for Economic Studies, Munich, Germany.

Feldstein, M. (1995), The Effect of Marginal Tax Rates on Taxable Income: A Panel Study of the 1986 Tax Reform Act. *Journal of Political Economy*, 103, 551–572.

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

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

Hausman, J.A. (1985), The Econometrics of Nonlinear Budget Sets. *Econometrica*, 53, 1255–1282.

Heckman, J.J. (1979), Sample Selection Bias as a Specification Error. *Econometrica*, 47, 153–161.

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

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

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

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

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

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

Lambert, P.J. and T.O. Thoresen (2009), Base Independence in the Analysis of Tax Policy Effects: with an Application to Norway 1992-2004. *International Tax and Public Finance*, 16, 219–252.

Lindsey, L. (1987), Estimating the Behavioral Responses of Taxpayers to Changes in Tax Rates: 1982-1984. *Journal of Public Economics*, 33, 173–206.

McFadden, D. (1984), “Econometric analysis of qualitative response models”. In Z. Griliches and M.D. Intriligator (eds.): *Handbook of Econometrics*, Vol. 2, Amsterdam: North-Holland, 1385–1457.

Moffitt, R. and M.O. Wilhelm (2000), "Taxation and the Labor Supply Decisions of the Affluent". In J. Slemrod (ed.): *Does Atlas Shrug? The Economic Consequences of Taxing the Rich*, New York: Russell Sage Foundation, 193–234.

Saez, E., J.B. Slemrod and S.H. Giertz (2009), The Elasticity of Taxable Income with Respect to Marginal Tax Rates: a Critical Review. NBER Working Paper No 15012, National Bureau of Economic Research, Cambridge, MA.

Saez, E., J.B. 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, 3–50.

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

Statistics Norway (2005), Labour Force Survey, 2004, http://www.ssb.no/aku_en/.

Statistics Norway (2006a), Income Statistics for Persons and Families, http://www.ssb.no/english/subjects/05/01/inntpf_en/.

Statistics Norway (2006b), Stock Statistics, http://www.ssb.no/english/subjects/11/01/aksjer_en/.

Statistics Norway (2006c), Wage Statistics. Wage Totals from End of the Year Certificate Register, http://www.ssb.no/english/subjects/06/05/lonnltreg_en/.

Sørensen, P.B. (2005), Neutral Taxation and Shareholder Income. *International Tax and Public Finance*, 12, 777–801.

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. *American Economic Review*, 96, 1384–1417.

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

Appendix A. Specification and estimation of discrete choice model

The discrete choice model presented in Section 2.1 is estimated for single females, singles 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 in the non-labor income variable.

To simplify the choice, we group the jobs into 11 categories based on weekly hours of work: $h_w \in \langle 0-5, 5-10, 10-15, ..., 45-50, 50+ \rangle$. As denoted in Section 2, 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 hours work $g_1(h)$. We assume that the densities of offered hours are uniform, except for peaks at long part-time (25–30) and full-time (35–40), whereas we let the individual differences be determined by the level of education.

The deterministic part of preferences is represented by the following "Box-Cox" type utility function, $v(C,h)=\alpha_0\frac{(C-C_0)^{\alpha_1}-1}{\alpha_1}+\gamma X\frac{(\bar{h}-h)^{\beta_1}-1}{\beta_1}$, where C_0 measures a minimum consumption level given by $60,000\sqrt{N}$ where *N* is the number of individuals in the household, $\bar{h}$ is defined as 80 hours per week and *h* is working hours per week such that $(\bar{h}-h)$ measures leisure per week, and *X* is a vector of taste-modifying variables. As seen in Section 2.1, Consumption is given by $C=f(hw,I)$. For couples, the income level of the spouse is assumed exogenously given and is included in non-labor income. We have accounted for that non-labor income may differ for different labor supply choices, as some transfer depends on this choice.

Information about actual and formal working time in main and secondary jobs and information on labor force status are obtained from the Labor Force Survey of 2004 (Statistics Norway, 2005). This is the main source of labor market statistics in Norway; about 24,000 individuals from representative selected families participate. Each respondent is asked about hours of work and attachment to the labor market in a reference week over eight subsequent quarters. Information about incomes, family composition, number of children, education, etc, are obtained from the Income Statistics for Persons and Families (Statistics Norway, 2006a) and merged with the Labor Force Survey by using a personal identification numbers. Based on information about labor force status from the Labor Force Survey, we have included wage earners and "potential" wage earners, coded as employed and home workers. Unemployed, self-employed, disabled and students are excluded from the sample. We restrict to persons aged between 25 and 62 and we define a person as non-participating if he or she works less than five hours per week.

Working time is measured as actual hours of work in both the main and second job, by using the average of reference week information for four quarters, provided by the Labor Force Survey. A key assumption is that this average gives a good proxy of a "normal" working week. An alternative to this measure of working hours is to use contractual hours of work. Not surprisingly, for two groups of wage earners with contractual hours of 37.5, the average yearly wage income for individuals with actual hours above 37.5 is considerable higher than for individuals working contractual hours or less. The reason is that individuals work overtime, and get paid for that through their standard wage or have the option to charge the employers for their extra workload. We deem that it is important to account for this characteristic of the labor market, also given that we focus on tax changes at high income levels in the present study.

If the respondent is only participating in the Labor Force Survey one quarter or if information on actual hours is missing (for example due to illness) then contractual hours is used instead. Contractual hours are also used if there is a large difference between contractual and actual hours, assuming that the latter may suffer more from measurement errors.

Since the Labor Force Survey does not contain any wage information, we computed hourly wage as yearly wage income (obtained from register-based tax return data) divided by annual hours per week (measured as 48 times average weekly hours). Then, the log of computed wage rates are regressed on individual characteristics, using 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), for females and a standard OLS regression for males[17]. Number of children and wealth are used as exclusion restrictions, under the hypothesis that these variables affect participation, but not hourly wage. We excluded individuals with improbable low or high computed hourly wage rates (under 60 or above 1,200 NOK in 2004) in the wage regression. A random effect is accounted for by adding an error term, based on a draws (30 draws per individual) from a normal distribution with standard errors according to the residuals in the wage regression.

Tables A1 and A2 report results of the wage equation regressions, whereas Tables A3 and A4 provide results of the labor supply model.

Table A1. Results of wage regressions for single males and males in couple.

Log(hourly wage)	Single males		Males in couple	
	Coefficient	Std error	Coefficient	Std error
Experience	0.0222***	(0.0023)	0.0271***	(0.0029)
Experience squared	-0.0004***	(0.0001)	-0.0005***	(0.0001)
Low education	-0.1121***	(0.0205)	-0.0847***	(0.0188)
High education	0.2313***	(0.0135)	0.2642***	(0.0128)
Residence in dense populated area	0.0834***	(0.0119)	0.1171***	(0.0127)
Non-west origin	-0.1264***	(0.0330)	-0.1530***	(0.0271)
Business code (ref. Public)				
Industry	0.1275***	(0.0151)	0.1634***	(0.0144)
Commerce	0.0241	(0.0160)	0.0877***	(0.0160)
Financial	0.1118***	(0.0179)	0.1560***	(0.0171)
Constant	4.8970***	(0.0262)	4.8678***	(0.0386)
Observations	3303		3 808	
R-square	0.169		0.194	

[17] The number of home working males are too small to account for selection effects

Table A2. Results of wage regressions for single females and females in couple, Heckman two-stage selection regression.

	Single females		Females in couple	
	Coefficient	Std error	Coefficient	Std error
Experience	0.0161***	(0.0020)	0.0124***	(0.0027)
Experience squared	-0.0003***	(0.0000)	-0.0002***	(0.0000)
Low education	-0.0651***	(0.0199)	-0.0835***	(0.0172)
High education	0.2027***	(0.0119)	0.2055***	(0.0115)
Residence in dense populated area	0.0650***	(0.0109)	0.0569***	(0.0107)
Non-west origin	-0.0183***	(0.0371)	-0.0373	(0.0252)
Business code (ref. Public)				
Industry	0.1379***	(0.0179)	0.1202***	(0.0159)
Commerce	-0.0031	(0.0132)	0.0072	(0.0126)
Financial	0.0953***	(0.0158)	0.1193***	(0.0147)
Constant	4.8695***	(0.0248)	4.9203***	(0.0394)
Participation				
Experience	0.0837***	(0.0236)	0.0782***	(0.0206)
Experience squared	-0.0017***	(0.0005)	-0.0017***	(0.0004)
Low education	-0.2429	(0.1874)	-0.2984*	(0.1236)
High education	0.5332***	(0.1386)	0.5548***	(0.1097)
Residence in dense populated area	0.0002	(0.1266)	0.0165	(0.0952)
Non-west origin	-1.0394***	(0.2213)	-0.7323***	(0.1396)
Business code (ref. Public)				
Industry	-0.1391	(0.1724)	0.2731	(0.1591)
Commerce	0.4291**	(0.1618)	0.1399	(0.1101)
Financial	-0.2086	(0.1539)	-0.4005***	(0.1051)
Number of children under 3 years	-0.3321*	(0.1373)	-0.2293	(0.1217)
Number of children under 6 years	-0.1817	(0.1559)	-0.1583	(0.1126)
Number of children under 12 years	-0.2481*	(0.1046)	-0.2210***	(0.0652)
Net wealth in 1000 NOK	-0.0021*	(0.0009)	-0.0012	(0.0009)
Constant	1.4425***	(0.2563)	1.3562***	(0.2998)
Mills lambda	-0.2147***	(0.0629)	-0.2208**	(0.0723)
Observations	3013		3927	
Censored observations	89		166	
Wald chi2	555.70		623.05	
Prob<chi2	0.0000		0.0000	

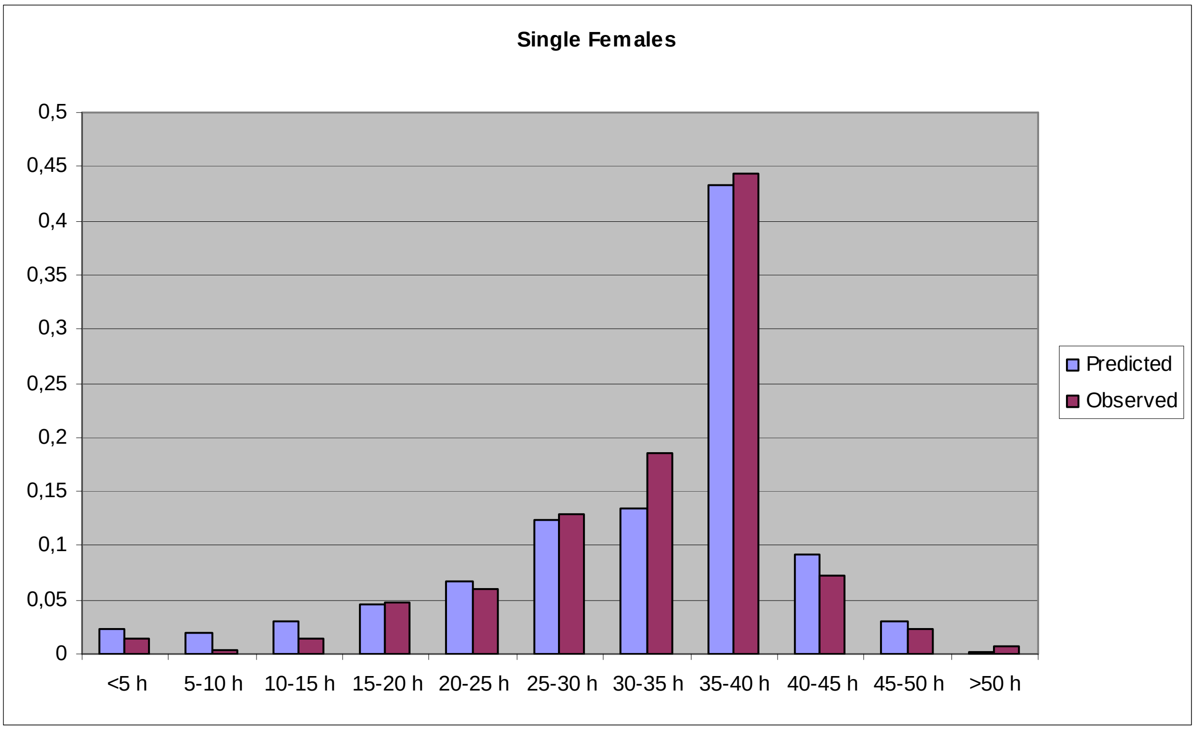
Table A3. Parameter estimates of the labor supply model, single females and single males

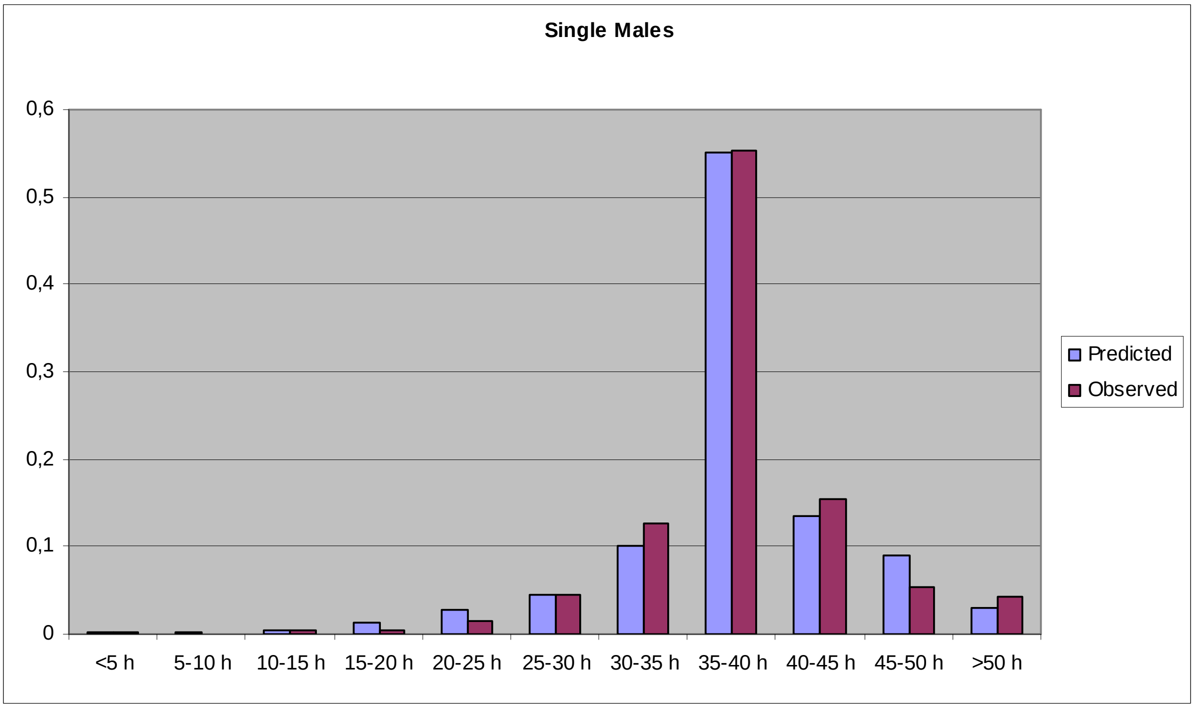
		Single females		Single males	
		Coefficient	Std Error	Coefficient	Std Error
Consumption					
Constant (Scale 10^{-4})	α_0	0.4591***	(0.0569)	0.9871***	(0.1487)
Exponent	α_1	0.8489***	(0.0626)	0.6090***	(0.0819)
Subsistence Level	C_o	$60{,}000\sqrt{N}$		$60{,}000\sqrt{N}$	
Leisure					
Age	γ_1	-0.0128	(0.0200)	-0.0310	(0.0447)
Age Squared	γ_2	0.0003	(0.0002)	0.0005	(0.0005)
High Education	γ_3	-0.0899	(0.0625)	-0.0014	(0.1047)
Low Education	γ_4	0.0136	(0.0865)	0.0104	(0.1257)
# Children under 6 years	γ_5	-0.1097*	(0.0548)	-0.2395*	(0.1005)
# Children above 6 years	γ_6	-0.0425	(0.0440)	-0.1465	(0.1036)
Residence in dense pop area	γ_8	-0.1427*	(0.0610)	0.0071	(0.0713)
Non-west Origin	γ_9	0.0697	(0.1405)	0.2313	(0.2668)
Constant	β_0	1.2579**	(0.4552)	2.2971*	(1.0347)
Exponent	β_1	-3.1311***	(0.3697)	-1.9795**	(0.6750)
Leisure x consumption	Δ	-0.0064	(0.0138)	-0.0189	(0.0824)
The parameters θ_F: $\log(\theta^{-1}) = f_{F1} + f_{F2}S$					
Constant	f_{F1}	1.6753	(0.9231)	0.5820	(3.3885)
Years of education	f_{F2}	-0.1292	(0.0769)	-0.0694	(0.2769)
Opportunity density of offered hours					
Part-time peak		0.1961**	(0.0629)	-0.2884**	(0.0947)
Full-time peak		1.1590***	(0.0486)	1.4055***	(0.0441)
Observations		3,036		3,356	

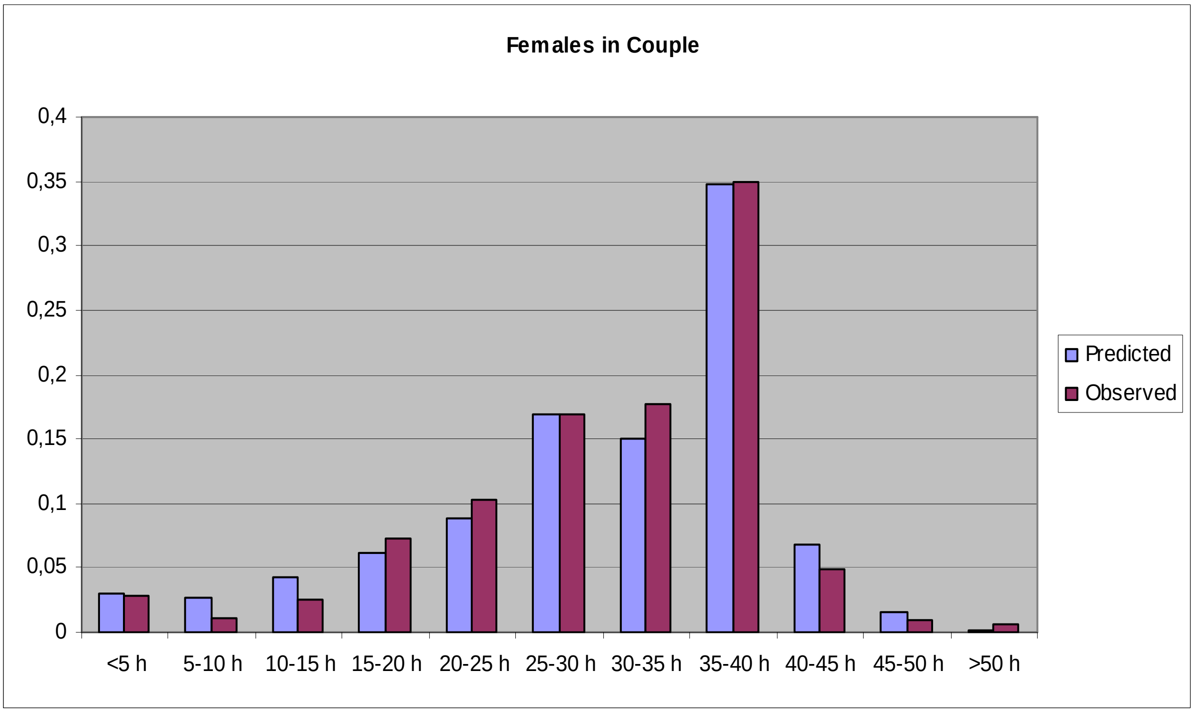
Table A4. Parameter estimates of the labor supply, females and males in couples

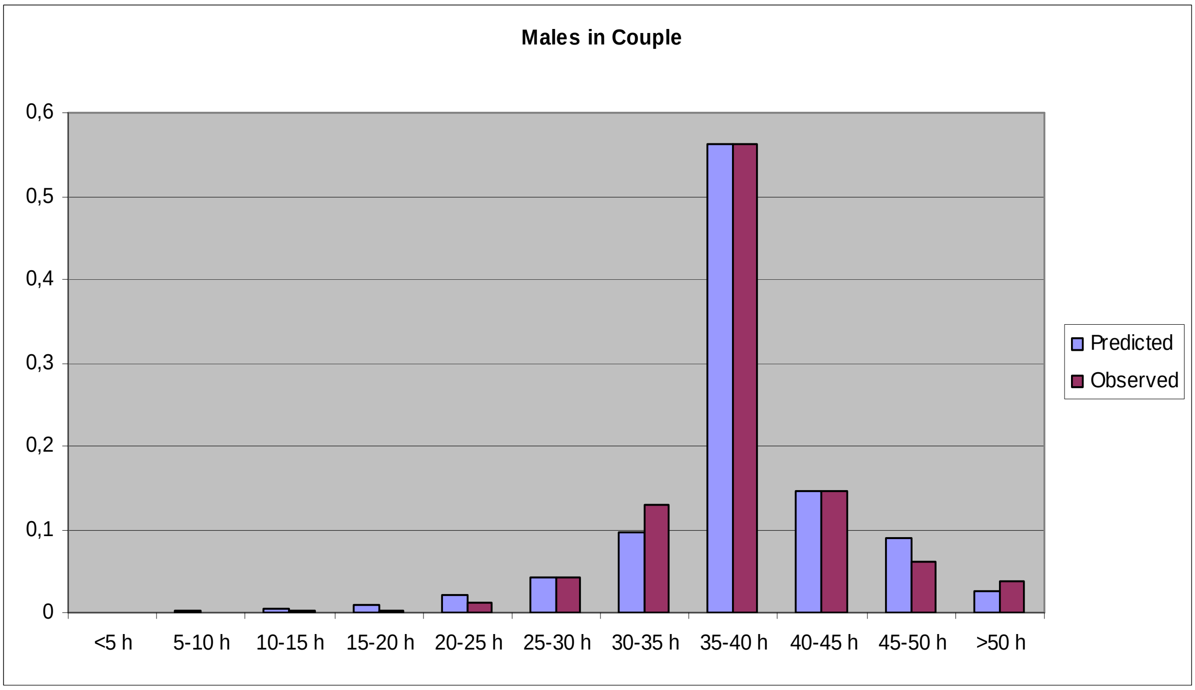
		Female in couple		Male in couple	
		Coefficient	Std Error	Coefficient	Std Error
Consumption					
Constant (Scale 10^{-4})	α_0	0.7267***	(0.0932)	2.3866***	(0.4197)
Exponent	α_1	0.8768***	(0.0416)	0.6058***	(0.0640)
Subsistence Level	C_o	$60{,}000\sqrt{N}$		$60{,}000\sqrt{N}$	
Leisure					
Age	γ_1	-0.0063	(0.0283)	-0.1147**	(0.0443)
Age squared	γ_2	0.0004	(0.0003)	0.0014**	(0.0005)
High education	γ_3	-0.2486***	(0.0543)	0.3163**	(0.1006)
Low education	γ_4	0.0466	(0.0964)	-0.3404**	(0.1207)
# Children under 6 years	γ_5	0.0253	(0.0517)	-0.2444**	(0.0802)
# Children above 6 years	γ_6	-0.0120	(0.0329)	-0.2931***	(0.0703)
Residence in dense pop	γ_7	-0.2607***	(0.0610)	0.0516	(0.0776)
Non-west origin	γ_8	0.1316	(0.1184)	0.1858	(0.1787)
Constant	β_0	1.4376*	(0.6439)	5.9957***	(1.2984)
Exponent	β_1	-2.7010***	(0.1771)	-1.9748***	(0.1850)
Leisure x consumption	Δ	0.0000	(0.0023)	-0.0980*	(0.0403)
The parameters θ_F: $\log(\theta^{-1}) = f_{F1} + f_{F2}S$					
Constant	f_{F1}	2.3445***	(0.5427)	-3.6852	(4.4697)
Years of education	f_{F2}	-0.1529**	(0.0471)	0.2181	(0.3156)
Opportunity density of offered hours					
Part-time peak		0.2828***	(0.0482)	-0.1491	(0.0909)
Full-time peak		0.9917***	(0.0450)	1.3829***	(0.0412)
Number of Observations		3,982		3,832	

In order to further evaluate the estimation results the figures below display the actual frequencies of working hours and the corresponding probability distribution, based on model simulations, for single females, single males, female and male in couple respectively. The simulated probabilities are derived by calculating the average probability for each hour choice, based on the individual probabilities.


Single Females
0,5
0,45
0,4
0,35
0,3
0,25
0,2
0,15
0,1
0,05
0
<5 h
5-10 h
10-15 h
15-20 h
20-25 h
25-30 h
30-35 h
35-40 h
40-45 h
45-50 h
>50 h
Predicted
Observed


Single Males
0,6
0,5
0,4
0,3
0,2
0,1
0
<5 h
5-10 h
10-15 h
15-20 h
20-25 h
25-30 h
30-35 h
35-40 h
40-45 h
45-50 h
>50 h
Predicted
Observed


Females in Couple
0,4
0,35
0,3
0,25
0,2
0,15
0,1
0,05
0
<5 h
5-10 h
10-15 h
15-20 h
20-25 h
25-30 h
30-35 h
35-40 h
40-45 h
45-50 h
>50 h
Predicted
Observed


Males in Couple
0,6
0,5
0,4
0,3
0,2
0,1
0
<5 h
5-10 h
10-15 h
15-20 h
20-25 h
25-30 h
30-35 h
35-40 h
40-45 h
45-50 h
>50 h
Predicted
Observed

Appendix B. Estimation of the elasticity of taxable income

A general specification

As discussed in Section 2, the standard framework for estimation of the elasticity of taxable income is to employ panel data information, typically estimating a model in differences for a 3-year span. In our case we utilize data over the time period 2000–2008. The individual specific effect is eliminated by first differencing. Other time invariant explanatory variables are added in as explanations to income growth.[18] Thus, letting Δ symbolize the differences we have that the difference in (log) taxable income is explained by differences in marginal tax rates, $1-\tau_i$, differences in virtual and non-labor income (included spouse income), R_i, and a set of socio-demographic variables, X_i:

$$
\text{(B.1)} \quad \begin{aligned}
\log\left(\frac{q_{it}}{q_{it-1}}\right) &= \kappa + \lambda_1 \log\left(\frac{1-\tau_{it}}{1-\tau_{it-1}}\right) + \lambda_2 \log\left(\frac{R_{it}}{R_{it-1}}\right) + X_i\omega + \xi_i \\
&= \kappa + \lambda_1 \Delta \log\left(1-\tau_i\right) + \lambda_2 \Delta \log R_i + X_i\omega + \xi_i .
\end{aligned}
$$

The key parameter is λ_1, which measures the uncompensated elasticity of taxable income. In the following we will present result for estimations of Equation (B.1), using panel data derived from administrative registers, with Income Statistics for Persons and Families as main source (Statistics Norway, 2006a). The income register contain information for the entire population in Norway (about 4.6 million in 2004). We will, however, restrict the sample to wage earners, defined as having wage income as their main source of income. We exclude individuals with positive income from self-employment or pensions. In addition, we restrict the sample to individuals with taxable labor income above percentile 33 (about 250 000 NOK in 2004) in the base year in our main analysis. As in the structural model we restrict the sample to individuals aged 25-62. We are left with about 5 million 3-year differences.

As the tax variables are endogenous, the main empirical strategy employs instrumental variable techniques (IV), and let the time span cover a tax reform to obtain exogenous variations in tax rates. Aarbu and Thoresen (2001) used the 1992 reform, here the 2006 reform is exploited. In order to let the exogenous variation determine the tax change, the standard procedure is to let the tax change be calculated on basis of first period income; see Auten and Carroll (1999), Moffitt and Wilhelm (2000), Gruber and Saez (2002),[19] and use an IV procedure to establish a predicted net-of-tax rate. In terms of two-stage least squares, in the first stage predicted net-of-tax rate changes are calculated by regressing

[18] Assuming that their relations to income change over time.

[19] Also similar to one of the two identification strategies used in Aarbu and Thoresen (2001).

the actual change, calculations of net-of-tax rates based on incomes in periods *t* and *t*-1, $\left[1-\tau_{it}\left(q_{it}\right)\right]/\left[1-\tau_{it-1}\left(q_{it-1}\right)\right]$, against the net-of-tax rate change instrument (the excluded variable) and all other explanatory variables, where the instrument is based on *t*-1 income inflated to the *t* level by using the actual wage growth over the period. The change in virtual and non-labor income, *R*, is instrumented by using the same exogenous tax rate change and by inflating non-labor income in the base year.

Definition of the dependent variable

In our study, the dependent variable is the growth in labor income which is the tax base for labor taxation, but not due to any deductions or exemptions. The elasticity we obtain can therefore be denoted as the elasticity of earned income. The effect from changes in the capital taxation is accounted for through including an income shifting control constructed by the exogenous change in marginal tax on capital interacted with the individual's capital income in the base year. Since we are restricting the set to wage earners, defined by that the main income comes from wage income, we simplify by assuming that other sources of income can be treated as exogenous.

The tax variables and other regressors

The tax reform which is used to generate exogenous variation in tax changes in the present analysis is described in Figure 1. We see that the reform changed maximum marginal tax rates from 55.3 percent in 2004 to 47.8 percent in 2007. We let these schedules decide marginal tax rates for actual taxation in 2004 and 2007: $\tau_{2004}\left(q_{2004}\right)$ and $\tau_{2007}\left(q_{2007}\right)$, and the "synthetic tax rate": $\hat{\tau}_{2007}((1+g)q_{2004})$, where g is an income growth factor. We will, however, use a slightly broader definition of marginal tax rates than what is presented in Figure 1, simulated by increasing the individuals' income by 5 percent and calculating the average tax rate on the added income. For identification reasons (see Section 2.2) it is preferable to have variation in tax rates which is not only directly dependent on base year taxable income. In that respect it is advantageous to have two different tax classes, separate or joint taxation of the couple, with different variation in tax rates. Moreover, it is advantageous that there is a separate net-of-tax rate schedule for people living in the northernmost areas of Norway (half of the county Troms and the county Finnmark).

The mean reversion control

The mean reversion problem has received extensive attention in the ETI-literature. In the ETI context the mean reversion problem refers to the way the instrumented tax variables are constructed: as first period income is used to define tax change instruments and tax reforms that are used to obtain tax change variation often imply either systematic reductions or increases in tax rates (predominantly focusing on changes in top marginal tax rates), the fact that individuals are temporary away from their permanent income path bring in systematic biases in estimates. For instance in terms of the Norwegian 2006 tax reform, 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*+1, and an income reduction will be recorded. Correspondingly, people with temporarily low income (non-treated) will be seen as increasing their income from *t*-1 to *t*, despite unaltered marginal tax rates. Thus, elasticity estimates will be negatively biased if not preventive measures are introduced.

To alleviate the mean reversion bias, Auten and Carroll (1999) suggest adding (log of) year *t*-1 income as a control variable. As shown by many analyses, Aarbu and Thoresen (2001) included, this control has a large influence on tax elasticity estimates. Gruber and Saez (2002) suggest an extension to the first period income control technique by also adding 10 splines defined in terms of first period income. We will also include centered polynomials of the base year income in some specifications which has been proposed as an alternative to splines. The main problem by 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, as the mean reversion control and the tax change instrument depend on the same variable (period *t*-1 income); see for instance Saez, Slemrod and Giertz (2010).[20] As already denoted, our empirical study benefits from other sources for variation than income alone (geography and different tax-classes), but we also include periods without tax reform in the data, in order to distinguish between mean reversion effects and tax responses.

Accounting for distributional trends

The spline function in the log of first period income is not only a control for (differentiated) mean reversion effects along the income scale, it may also be seen as accounting for evolutions of income distributions. For example a trend towards increasing inequality may give spurious correlation between lowered tax rates for high income individuals and their growth rates. For example the large elasticities measured by the tax reforms in the 80s in the US, was most likely influenced by the lack of control for the increasing inequality, which happened for other reasons over the same period.

Controlling for trends in income distributions is expected to be less important in the case of Norway, as we mainly focus on effects on earned income. Capital income has been the main contributor to the increased income inequality after the 1992 reform, see for instance Lambert and Thoresen (2009).

[20] A collinearity problem would emerge, leading into less robust estimates.

The danger of income distribution evolvements biasing the elasticity estimates is also reduced by controlling for a large number of other individual characteristics. We have had access to a number of socio-demographic characteristics, such as age, years of education, type of education, marital status, number of children, geographical location and area of origin. Hypotheses of how these characteristics affect income growth can be made: for instance, we expect the presence of young children to limit income growth and education length to have a positive effect.

Estimation results

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

We find the specifications (3)–(6) most convincing, as we believe it is not sufficient to include a linear control for the mean reversion.[21] We see that the results are less influenced by using either splines or polynomials as control variables.

[21] When including only a linear control for mean reversion, the estimated elasticity becomes very dependent on sample restrictions.

Table B1. Estimation results: wage income growth as the dependent variable

Dependent variable:	Mean Reversion Control					
	Log base year income		10 Splines of base year income		3 degree polynomial of base year income	
	(1)	(2)	(3)	(4)	(5)	(6)
Net-of-tax 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)
Non-labor 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)
Newborn	-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 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 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-west origin	-0.0431***	-0.0420***	-0.0432***	-0.0421***	-0.0432***	-0.0421***
	(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)
Dense 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)
Occupation dummies	Yes	Yes	Yes	Yes	Yes	Yes
$\log(\text{income}_{it}/\text{median}_t)$	-0.1124***	-0.1055***				
	(0.0004)	(0.0004)				
10 linear Splines			Yes	Yes		
3 polynomial					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	4,933,291	4,331,276	4,933,291	4,331,276	4,933,291	4,331,276

Table B.2 provide the full regression output when categorizing the sample with respect to gender and civil status. In all regressions we have included a centered third degree polynomial of the base year income as mean reversion control. Results show less variation with respect to these categorizations.

Table B2. Estimation results: wage income growth as the dependent variable by couple and gender

	Female, single	Male, single	Female, couple	Male, couple
Net-of-tax rate elasticity	0.0340***	0.0298***	0.0410***	0.0448***
	(0.0061)	(0.0059)	(0.0049)	(0.0031)
Income shifting control	0.0112***	0.0120***	0.0089***	0.0112***
	(0.0005)	(0.0005)	(0.0004)	(0.0002)
Wealth	-0.0005***	-0.0001*	-0.0008***	0.0000
	(0.0001)	(0.0000)	(0.0000)	(0.0000)
Age	0.0002	-0.0027***	0.0074***	-0.0014***
	(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	0.0178***	0.0067***	0.0136***	0.0124***
	(0.0011)	(0.0010)	(0.0005)	(0.0004)
Newborn	-0.1399***	-0.0073	-0.1430***	-0.0137***
	(0.0026)	(0.0040)	(0.0008)	(0.0005)
No. children under 6	0.0383***	-0.0021	0.0371***	0.0063***
	(0.0018)	(0.0026)	(0.0006)	(0.0003)
No. children above 6	0.0147***	0.0123***	0.0093***	0.0037***
	(0.0005)	(0.0008)	(0.0003)	(0.0002)
Non-west origin	-0.0326***	-0.0573***	-0.0272***	-0.0510***
	(0.0019)	(0.0020)	(0.0013)	(0.0009)
Residence in Oslo	0.0076***	-0.0071***	0.0108***	0.0008*
	(0.0006)	(0.0006)	(0.0005)	(0.0003)
Dense populated area	0.0120***	0.0058***	0.0120***	0.0087***
	(0.0009)	(0.0007)	(0.0006)	(0.0004)
Years of education	0.0139***	0.0159***	0.0156***	0.0120***
	(0.0002)	(0.0001)	(0.0001)	(0.0001)
Occupation dummies	Yes	Yes	Yes	Yes
3 polynomial	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
Constant	-0.0154*	0.0672***	-0.2057***	0.1363***
	(0.0075)	(0.0071)	(0.0063)	(0.0042)
Number of observations	576,232	959,151	1,109,651	2,287,960

Note: 3 degree polynomial of base year income is used as mean reversion control

Robustness checks

Here we present robustness test for the sample cut-off value and for alternative time spans, since the choice of including individuals above percentile 33 and the choice of 3 year panels both are rather ad-hoc. In table B3 we present results for the net-of-tax rate elasticity for alternative cut-off rules.[22] In regression 1 we include all individuals in percentile 25 or above and in regression 3 we include all in percentile 40 or above. We expect that the estimate of net-of-tax rate elasticity is independent of this choice, as individuals in the control group, independent of the cut-off point, were not affected by the reform. The results uncover that there are very small differences in the net-of-tax rate with respect to the choice of cut-off value.

Table B3. Robustness checks: Cut-off value

	(1) >percentile 25		(2) >percentile 33		(3) >percentile 40	
	Coeff.	Std error	Coeff.	Std error	Coeff.	Std error
Net-of-tax elasticity	0.0520***	(0.0023)	0.0531***	(0.0023)	0.0534***	(0.0022)
Number of observations	5,486,168		4,933,291		4,439,785	

The three year span has been proposed in the literature to allow for some time for individuals to respond to tax changes. Still, the choice is ad-hoc and here we present the results for alternative spans, 1 to 4 years. The regressions include no income effect and again the third degree polynomial is used as mean reversion control and the cut-off is chosen at percentile 33. The results are relatively robust to alternative spans with lowest elasticity of 0.032 for 1 year differences. The likely reason for that is that wage earners do not respond immediately to tax changes. The elasticity seems to be highest using 3-year spans.

Table B4. Robustness checks: Alternative time spans

	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

[22] Table B3 and B4 show results for the net-of tax rate elasticity only, but these results are based the same specification as in Table B2 without virtual income control.