Statistisk sentralbyrå
Statistics Norway

Substitution and income effects of labor income taxation

Michael Graber, Morten Håvarstein, Magne Mogstad, Gaute Torsvik and Ola L. Vestad

TALL SOM FORTELLER

DISCUSSION PAPERS

1037

Discussion Papers: comprise research papers intended for international journals or books. A preprint of a Discussion Paper may be longer and more elaborate than a standard journal article, as it may include intermediate calculations and background material etc.

The Discussion Papers series presents results from ongoing research projects and other research and analysis by SSB staff. The views and conclusions in this document are those of the authors.

Published: April 2026

Abstracts with downloadable Discussion Papers
in PDF are available on the Internet:

https://www.ssb.no/discussion-papers
http://ideas.repec.org/s/ssb/dispap.html

ISSN 1892-753X (electronic)

Abstract

The elasticity of taxable income (ETI) parameter is a key quantity in empirical analysis of tax policy and labor supply. We examine when a commonly applied class of ETI estimands can be used to learn about individuals' ETI parameters and their (un)compensated elasticities of labor supply. We begin by providing necessary and sufficient conditions for these estimands to be given a causal interpretation as a positively weighted average of heterogeneous ETI parameters. We then apply these results to empirically analyze a reform of the Norwegian tax system that reduced the marginal tax rates on middle and high incomes. The estimated ETI parameters increase steadily with income, meaning high-income individuals are more responsive to tax changes than middle-income individuals. Next, we show how (un)compensated elasticities of labor supply can be bounded directly from the ETI estimands, or point identified by combining these estimands with estimates of earnings responses to lottery winnings. The results suggest an (un)compensated elasticity of 0.1 (0.0) for middle-income individuals. The (un)compensated elasticity estimates increase steadily with income to around 0.45 (0.3) for high-income individuals. These findings imply a substantial excess burden of taxation, and that reducing top-income tax rates would increase tax revenue. Our findings are also informative about how the intertemporal elasticity of substitution and the Frisch elasticity vary across the income distribution.

Keywords: compensated elasticity, elasticity of intertemporal substitution, elasticity of taxable income, Frisch elasticity, income effects, labor supply, substitution effect.

JEL classification: C20, D15, H20, J22

Acknowledgements: Utkarsh Dandanayak provided outstanding research assistance. Håvarstein, Mogstad, and Torsvik gratefully acknowledge financial support from the Norwegian Research Council. We thank Rolf Aaberge, Deniz Dutz, Mikhail Golosov, Kristen Vamsæter, and Harald Uhlig for valuable comments.

Address: Michael Graber, Statistics Norway. E-mail: michael.graber@ssb.no; Morten Håvarstein, University of Chicago and University of Oslo. E-mail: morten.havarstein@econ.uio.no; Magne Mogstad, University of Chicago and Statistics Norway. E-mail: magne.mogstad@gmail.com; Gaute Torsvik, University of Oslo. E-mail: gaute.torsvik@econ.uio.no; Ola L. Vestad, Statistics Norway. E-mail: ola.vestad@ssb.no.

Sammendrag

Elastisiteten til skattepliktig inntekt (elasticity of taxable income, ETI) måler hvordan skattepliktig inntekt påvirkes av endringer i marginalskattesatser og er sentral i empiriske analyser av skattepolitikk og arbeidstilbud. I denne artikkelen undersøker vi når mye brukte ETI-estimatorer kan gi informasjon om individers ETI-parametere og deres (u)kompenserte arbeidstilbudselastisiteter. Vi utleder nødvendige og tilstrekkelige betingelser for at disse estimatorene kan tolkes kausalt som et positivt vektet gjennomsnitt av heterogene ETI-parametere. Vi anvender deretter resultatene i en empirisk analyse av en norsk skattereform, 2006-reformen, som reduserte marginalskatten på middels høye og høye inntekter. Estimatene viser at ETI øker med inntekt, det vil si at høyinntektsgrupper responderer sterkere på endringer i marginalskattesatser enn grupper med middels inntekt. Videre viser vi hvordan (u)kompenserte arbeidstilbudselastisiteter kan avledes fra ETI-estimatene ved å etablere øvre og nedre grenser, og hvordan punktidentifikasjon kan oppnås ved å kombinere ETI-estimatene med estimater av inntektsresponser på lotterigevinster. Resultatene viser at de (u)kompenserte elastisitetene øker med inntekt, fra om lag 0,1 (0,0) for personer med middels inntekt til om lag 0,45 (0,3) for personer med høye inntekter. Funnene impliserer et betydelig effektivitetstap ved beskatning av arbeidsinntekt, og at lavere toppskattesatser kan øke skatteinntektene. Resultatene gir også innsikt i hvordan den intertemporale substitusjonselastisiteten og Frisch-elastisiteten varierer med inntekt.

1 Introduction

The elasticity of taxable income (ETI) parameter measures how taxable income responds to reforms that change the marginal tax rate. It is a key quantity in tax policy for assessing how exogenous changes in tax rates will causally affect income and tax revenue (Auten and Carroll, 1999). Also, it is often interpreted as a compensated elasticity of labor supply (Saez et al., 2012), which can be used to assess the excess burden of taxation (Feldstein, 1999). The goal of our paper is to show when and how a commonly applied class of ETI estimands can be used to learn about individuals' ETI parameters and their (un)compensated elasticities of labor supply.

In Section 2, we provide necessary and sufficient conditions for the ETI estimands to be given a causal interpretation as a positively weighted average of heterogeneous ETI parameters. This identification result is constructive, leading to empirical specifications and estimators that can be easily implemented. These specifications differ from the ETI estimands commonly used, which fail to satisfy the conditions for a causal interpretation when elasticities are heterogeneous. A causal interpretation with heterogeneous elasticities requires both a specification that controls for past income sufficiently flexibly, and no comparison of earnings responses between individuals who experience changes in marginal tax rates of varying degrees.

The identification results in Section 2 guide our empirical analysis in Section 3 of a reform of the Norwegian tax system that reduced the marginal tax rates on middle and high incomes. We find that commonly used ETI estimands that are causal only under constant elasticities understate the average ETI parameter. The specifications that can be given a causal interpretation produce estimates of an average ETI parameter across the income distribution of around 0.21. This average misses a great deal: the estimated ETI parameters increase steadily with income, meaning high-income individuals are more responsive to tax changes than middle-income individuals. For example, the ETI parameter at the median income is less than 0.1, while it is 0.35 at the 90th percentile.

In Section 4, we show how (un)compensated elasticities of labor supply can be point or partially identified from our preferred causal specification of the ETI estimands. The continuous choice labor supply model we consider allows for income effects, elasticities that vary across the income distribution, and correlation between individuals' productivity and their unobserved taste for work.[1] This model allows us to express the ETI estimands in terms of the (un)compensated elasticities plus a bias term that would be observable in the data if income effects were known or could be estimated. This observation motivates and guides the analysis in the remainder of the paper, where we consider different approaches to point or partially identify the labor supply elasticities and functionals of these elasticities, such as the excess burden of taxation.

The first approach we consider is to construct bounds using the Engel aggregation condition, which implies that income effects are bounded between -1 and 0. The bounds suggest that the compensated and uncompensated elasticities of high-income individuals at the 90th percentile are at least 0.35 and 0.1, respectively. By comparison, the compensated elasticity of middle-income individuals close to the median income is bounded between 0.05 and 0.5, while their uncompen-

[1] Our model nests typical labor supply models in public finance (e.g., Feldstein, 1999; Gruber and Saez, 2002 and Saez et al., 2012). An alternative class of models represents the budget with a discrete set of alternatives and assumes that individuals' productivity is independent of their unobserved taste for work. This taste also follows a known parametric distribution. See Aaberge and Colombino (2018) and the references therein for discussion of these models and the findings they produce.

sated elasticities are close to zero or negative. These bounds can be tightened considerably by ruling out implausibly small and large income effects. Both the upper bound on the compensated elasticities and the lower bound on the uncompensated elasticities then become more informative. For example, if income effects are assumed to lie between –0.05 and –0.25, then the compensated elasticity of middle-income individuals must lie between 0.05 and 0.15, whereas the compensated elasticity of high-income individuals is bounded between 0.4 and 0.5.

There are two ways to move from partial to point identification. One is to make stronger assumptions. Gruber and Saez (2002) assume that income and substitution effects do not vary across the income distribution. The (un)compensated elasticity is then point identified. We test this assumption and find that it is strongly rejected by the data.

Another possible way to achieve point identification is to use more data. We pursue such an approach in Section 5, combining the ETI estimands with external estimates of income effects. A number of studies have estimated income effects from lottery winnings that create plausibly exogenous variation in unearned income, holding fixed all other determinants of behavior, such as preferences and wages.[2] We follow this approach and use Norwegian data on lottery winnings to estimate earnings and employment responses to exogenous changes in unearned income. We show how these responses allow us to infer the income effects on the intensive margin that we need to point identify the (un)compensated elasticities. The resulting point estimates suggest an (un)compensated elasticity of 0.1 (0) for middle-income individuals at the median income. The (un)compensated elasticity estimates increase steadily with income to around 0.45 (0.3) for high-income individuals at the 90th percentile.

In Section 6, we show that the findings in Section 5 imply a substantial excess burden of taxation, and that reducing the top-income tax rate would significantly increase tax revenue. We find that these conclusions contrast sharply with the results we obtain from conventional calculations that use the common ETI estimand, ignore income effects, and assume constant labor supply elasticities. The excess burden then becomes much lower, and the revenue-maximizing top-income tax rate moves closer to the observed one.

In Section 7, we consider the implications of our findings for other parameters of interest, including the elasticity of intertemporal substitution (EIS) of consumption and the Frisch elasticity. The EIS determines how agents adjust consumption in response to intertemporal price changes, while the Frisch elasticity measures the response of labor supply to predictable wage changes. We estimate a Frisch elasticity (EIS) of 0.1 (0.5) for middle-income individuals at the median income. The Frisch elasticity (EIS) estimates increase steadily with income to around 0.5 (2.5) for high-income individuals at the 90th percentile.

Our paper contributes to a large set of studies that have used models of labor supply to try to recover income and substitution effects from observational variation in unearned income, wages, and tax rates. The models, data, and findings have been summarized and critiqued in multiple review articles, including Blundell and Macurdy (1999), Keane (2011), Killingsworth and Heckman (1986), Pencavel (1986), and Saez et al. (2012). As emphasized in these reviews, there is no consensus about the size of income and substitution effects and how they vary across the income distribution. As a result, it is difficult to draw credible conclusions about parameters that depend on these income and substitution effects, such as the excess burden of taxation and the revenue-

[2]See, for example, Cesarini et al. (2017), Golosov et al. (2024), and Imbens et al. (2001).

maximizing tax rates.

A key reason for the lack of consensus is that it has been difficult to separately identify income and substitution effects without strong assumptions on functional form and the distribution of unobservables. Our paper shows how to construct informative bounds on (un)compensated elasticities from variation in take-home pay that arises from tax reforms. We also show how income and substitution effects can be point identified by combining the variation in take-home pay that arises from tax reforms with plausibly exogenous changes in unearned income. By doing so, our paper offers credible evidence on income and substitution effects, how they vary across the income distribution, and their implications for the excess burden of taxation and the revenue-maximizing top-income tax rate.

Our paper also contributes to the ETI literature, which analyzes how taxable income responds to reforms that change the marginal tax rates.[3] We extend the usual identification argument to allow for heterogeneity in the individuals' ETI parameters. Our necessary and sufficient conditions offer a blueprint for estimating ETIs while allowing for heterogeneous elasticities. The variation in elasticities across the income distribution that we find is significant, both statistically and economically. An average elasticity is therefore far from a sufficient statistic to predict how counterfactual changes in marginal tax rates would affect aggregate earnings and tax revenue.

Section 2 of our paper is related to Kumar and Liang (2020), which considers the causal interpretation of the ETI estimand without covariates, under the assumption that the tax system is randomly assigned across individuals. They show that this estimand is generally not equal to a *particular* weighted average of individual ETI parameters, unless the instrument is valid and the individual ETI parameters are homogeneous.

Our results about the ETI estimand differ in several ways. First, we consider the causal interpretation of the ETI estimand with covariates and do not assume that the tax system is randomly assigned. The inclusion of covariates in the theoretical results is important, since empirical work tries to flexibly control for individual characteristics such as past income. Second, we provide sufficient and necessary conditions for the ETI estimand to recover *any* positively weighted average of individual ETI parameters. This is arguably a minimal requirement for the ETI estimand to be an interesting quantity, but it is not sufficient. We therefore strengthen our result by showing *the* causal interpretation of the ETI estimand as a specific positively weighted average of individual ETI parameters. Third, we consider the bias of the ETI estimand due to income effects and how it can be corrected for by constructing bounds, invoking auxiliary assumptions, or using additional data.

Our paper is also related to a set of empirical studies that have used lottery winnings to estimate wealth and income effects (Bulman et al., 2021; Cesarini et al., 2017; Golosov et al., 2024; Imbens et al., 2001, and Picchio et al., 2018). To measure how lottery winnings are allocated over time, studies often rely on strong assumptions about intertemporal preferences or constraints. In contrast, our rich administrative data allow for imputing consumption and savings over time (Eika et al., 2020), which means we can measure unearned income in each period without relying on additional assumptions. Thus, our paper offers credible evidence on income effects and how they vary across the income distribution without relying on assumptions about how households

[3]Notable contributions to this literature include Auten and Carroll (1999), Burns and Ziliak (2017), Feldstein (1995), Gruber and Saez (2002), Kleven and Schultz (2014), and Kopczuk (2005). See Saez et al. (2012) for a review and Neisser (2021) for a meta-analysis.

allocate their wealth over time.

Lastly, our paper provides new evidence on the Frisch elasticity and the EIS, and how these elasticities vary across the income distribution. The Frisch elasticity plays a key role in the design of tax policy (see, e.g., Conesa et al., 2009) and in business cycle analysis (see, e.g. Prescott, 2006). The EIS is relevant both for understanding consumption and saving behavior (Hall, 1988) and for the design of long-run capital taxation (Straub and Werning, 2020). Because of their importance, a large body of work attempts to estimate the Frisch elasticity (see, e.g., MaCurdy, 1981 and Altonji, 1986) and the EIS (see, e.g., Hansen and Singleton, 1983 and Attanasio and Weber, 1995). Keane and Neal (2025) review the literature on Frisch elasticities, arguing that estimates vary considerably across studies. Havránek (2015) reviews the EIS literature, concluding that estimates vary greatly across studies. An important question is what explains the differences in findings. A number of possible explanations have been discussed (see, e.g., Attanasio et al., 2018; Browning and Lusardi, 1996, and Chetty et al., 2011). We contribute by showing that both the EIS and the Frisch elasticity vary considerably due to population heterogeneity in the earnings responses at the intensive margin. The elasticities are small but significant for middle-income individuals, and fairly large and significant for individuals with higher incomes.

2 Interpretation of ETI estimands

We now introduce and analyze a class of commonly used ETI estimands. We first describe the research design and data that form the basis for these estimands. We then discuss the causal interpretation of these estimands, providing necessary and sufficient conditions for causality. Although our discussion centers on the Norwegian tax reform and data, the arguments can easily be applied to other settings and reforms. The details of the Norwegian reform and data are discussed in Section 3.

2.1 Research design

Figure 1 illustrates the research design and the data that it will use. The figure considers a tax system that changed from T_0 to T_1 in 2006, decreasing the marginal tax rates on incomes above $\bar{Y}$. The reform cohorts ($G = 1$), which consist of observations from 2003-2007, experienced the *actual* reform in 2006. The placebo cohorts ($G = 0$), consisting of observations from 1998-2002, experience no change in the tax system. For each cohort type, we divide the data into a pre-period consisting of the first three years (event time $t = -3, -2, -1$) and a post-period covering the last two years ($t = 0, 1$).

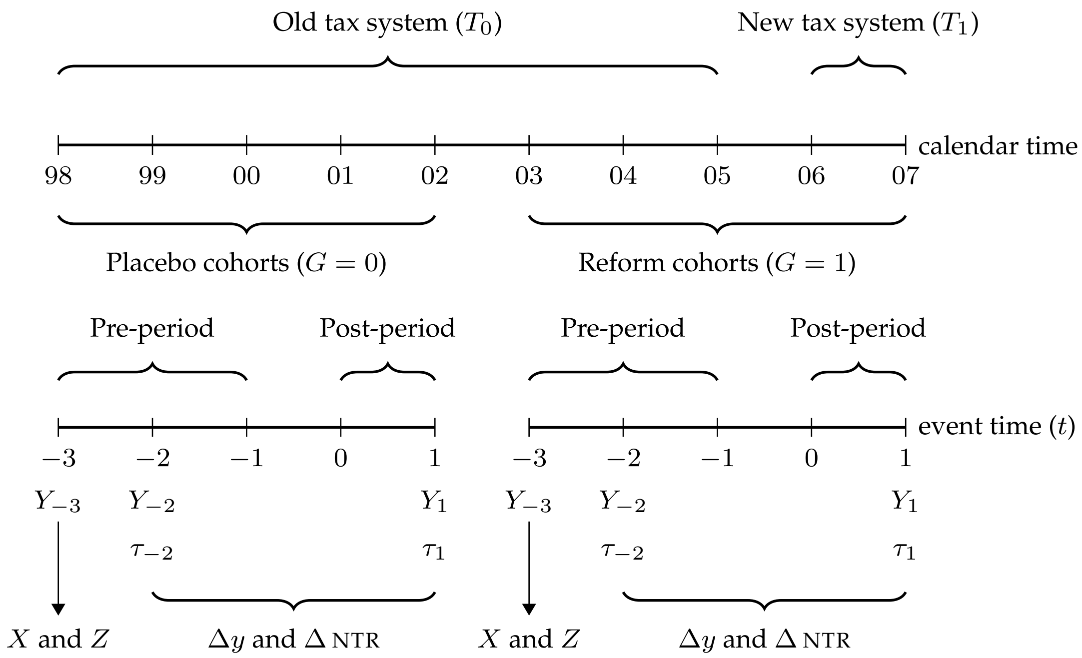
We observe the earnings Y and marginal tax rates τ for every individual at each event time:

$$\tau_t = T_0'(Y_t) + \mathbb{1}[t \geq 0]G\Big(T_1'(Y_t) - T_0'(Y_t)\Big),$$

where T_d' denotes the derivative of tax function T_d. Thus, the marginal tax rate of an individual depends on her earnings Y and whether she faces the old (T_0) or the reformed tax system (T_1).

To address the simultaneity between marginal tax rates and earnings, the empirical ETI literature uses *simulated instruments*, defined as the predicted percentage change in net-of-tax rates

Figure 1: Anatomy of the research design



Notes: The figure presents the research design, illustrates the timing of the reform, and introduces notation.

because of the reform:

$$Z \equiv \log\left(\frac{1 - T_1'(X)}{1 - T_0'(X)}\right), \tag{1}$$

where X is initial income, defined as earnings at event time -3.

The variable Z can take more than two values. It captures not only whether the reform changes the marginal tax rates of some (treatment group with $Z > 0$), but not all individuals (control group with $Z = 0$), but also that the magnitude of the change may differ across earnings levels. Hence, treatment intensity can vary across treated individuals.

2.2 ETI estimands and regression model

A possible estimand for the elasticity of taxable income (ETI) parameter is the difference-in-differences estimand that compares the earnings- and marginal tax rate growth of the treatment group to that of the control group:

$$\beta^{\text{DD}} \equiv \frac{\overbrace{\mathbb{E}[\Delta y \mid G=1, Z>0] - \mathbb{E}[\Delta y \mid G=1, Z=0]}^{\equiv RF \text{ (Earnings DiD)}}}{\underbrace{\mathbb{E}[\Delta\,\text{NTR} \mid G=1, Z>0] - \mathbb{E}[\Delta\,\text{NTR} \mid G=1, Z=0]}_{\equiv FS \text{ (Net-of-tax rate DiD)}}},$$

where $\Delta y \equiv \log Y_1 - \log Y_{-2}$, and $\Delta\,\text{NTR} \equiv \log(1-\tau_1) - \log(1-\tau_{-2})$.

A concern with this estimand is that because the reform changed marginal tax rates on incomes above $\bar{Y}$, the treatment group has higher initial incomes than the control group. Thus, if individual income growth depends on initial income, either due to mean reversion or differential underlying income growth, parallel trends are unlikely to hold (Weber, 2014). To address this concern, the literature uses the placebo cohorts to estimate the difference in earnings and net-of-tax rate growth between the treatment and control groups over the period when no tax reform occurred. The triple difference estimand β^{DDD} subtracts these placebo differences from the reform ones:

$$\beta^{\text{DDD}} \equiv \frac{\overbrace{RF}^{\text{Earnings DiD}} - \overbrace{\big(\mathbb{E}[\Delta y \mid G=0,\, Z>0] - \mathbb{E}[\Delta y \mid G=0,\, Z=0]\big)}^{\text{Placebo earnings DiD}}}{\underbrace{FS}_{\text{Net-of-tax rate DiD}} - \underbrace{\big(\mathbb{E}[\Delta\,\text{NTR} \mid G=0,\, Z>0] - \mathbb{E}[\Delta\,\text{NTR} \mid G=0,\, Z=0]\big)}_{\text{Placebo net-of-tax rate DiD}}}. \tag{2}$$

The triple difference estimand in (2) is nested by the following two-stage least squares (TSLS) regression model:

$$\Delta y = \alpha_0^y G + \beta \Delta\,\text{NTR} + f(X; \alpha^y) + u^y, \tag{3}$$

$$\Delta\,\text{NTR} = \alpha_0^{\text{NTR}} G + \alpha G h(Z) + f(X; \alpha^{\text{NTR}}) + u^{\text{NTR}}, \tag{4}$$

where h is a monotonic function of the simulated instrument that is strictly increasing at zero, and f satisfies

$$f(X; \alpha) = \sum_{j=1}^{J} \alpha_j m_j(X)$$

for a collection of non-linear functions $m_1, .., m_J$. We refer to the coefficient β as the *elasticity of taxable income* (ETI) estimand.

The TSLS model nests the triple differences model in equation (2) when $h(Z)$ is binary (equal to $\mathbb{1}[Z > 0]$) and f is a constant plus the indicator variable $\mathbb{1}[X \geq \bar{Y}]$. However, the TSLS model allows the researcher to choose a rich specification of f, thereby controlling flexibly for differential earnings- and tax rate growth, and to use variation in the intensity of treatment when the simulated instrument takes more than two values. The empirical ETI literature typically chooses $h(Z) = Z$ and specifies f to be a polynomial or spline function of X (see, e.g., Auten and Carroll (1999), Gruber and Saez (2002), and Kleven and Schultz (2014)).

2.3 Potential earnings model

To consider the causal interpretation of the ETI estimand (defined by equations (3) and (4)), it is necessary to introduce a potential earnings model that links the ETI estimand and the data to individual ETI parameters.

Given the tax systems T_0 and T_1, we let $Y_t(d)$ denote period t potential earnings under tax system T_d. Similarly, $\tau_t(d) \equiv T_d'(Y_t(d))$ denotes their potential marginal tax rate and $\text{NTR}_t(d) \equiv \log(1 - T_d'(Y(d)))$ their potential log net-of-tax rate. The potential outcomes map to observed outcomes through,

$$Y_t = Y_t(0) + \mathbb{1}[t \geq 0]\mathbb{1}[Z > 0]G\left[Y_t(1) - Y_t(0)\right], \tag{5}$$

$$\text{NTR}_t = \text{NTR}_t(0) + \mathbb{1}[t \geq 0]\mathbb{1}[Z > 0]G\left[\text{NTR}_t(1) - \text{NTR}_t(0)\right]. \tag{6}$$

Following Saez et al. (2012), we assume the potential earnings function is,

$$\log Y_t(\text{NTR}, d) = \zeta \times \text{NTR} + \nu_t(d), \tag{7}$$

where the parameters $\nu_t(0)$ and $\nu_t(1)$ can vary freely across individuals and ζ is the individual's ETI parameter.[4] Unless otherwise noted, we also allow ζ to vary freely across individuals. The specification in (7) allows the tax system to affect earnings both through marginal tax rates NTR and other channels ν.

It is useful to consider a set of assumptions commonly invoked to give instrumental variable estimands a causal interpretation, which imposes additional restrictions on the potential outcomes. Throughout the paper, we impose the following common trends assumption to recover the "reduced form" effects of the reform on earnings and the "first stage" effects of the reform on the marginal tax rates:

Assumption 1 (Common trends). *In the absence of the tax reform, average changes in earnings and marginal tax rates satisfy*

$$\mathbb{E}\left[\Delta \log Y(0) \mid G, X\right] = \kappa^y G + f^y(X), \tag{8}$$

$$\mathbb{E}\left[\Delta\, \text{NTR}(0) \mid G, X\right] = \kappa^{\text{NTR}} G + f^{\text{NTR}}(X), \tag{9}$$

for arbitrary functions f^y and f^{NTR}.

The restriction in Assumption 1 is that there is no interaction between G and X in the average growth in earnings (or marginal tax rates) in the absence of the reform. It means that the aggregate growth due to calendar-time effects (for example, due to business cycles) is allowed to vary freely over time, and that any idiosyncratic growth can vary freely across individuals depending on their initial income X.

Our second assumption is an exclusion restriction, which implies that individual earnings are affected by the tax reform only through its effects on marginal tax rates.

Assumption 2 (Exclusion restriction). $\nu_t(0) = \nu_t(1)$ *with probability one.*

As shown below, this restriction implies no income effects, an assumption we will relax in Section 4.

Lastly, we consider the assumption that the tax reform weakly decreases the marginal tax rates for all:

Assumption 3 (Monotonicity). $\text{NTR}(1) \geq \text{NTR}(0)$ *with probability one.*

This assumption is common in the program evaluation literature when allowing for treatment effect heterogeneity. It is typically necessary to ensure that standard IV estimands reflect positively weighted averages of individual treatment effects: see, e.g., Imbens and Angrist (1994).

[4]The log specification implicitly assumes no extensive margin responses to the tax reform. In Section 3, we test this assumption empirically and find that the reform we consider only had a limited effect on the extensive margin employment decision. To understand how these small extensive margin responses could bias our results, Appendix C develops bounds on the ETI while allowing for extensive margin responses. The estimated bounds are narrow and centered around the point estimates.

2.4 Necessary and sufficient conditions for the ETI estimand to be causal

We now provide a characterization of the ETI estimand in terms of group-specific average ETIs ζ under different choices of f and h. Denote by $\Phi \equiv \text{NTR}(1) - \text{NTR}(0)$ the effect the tax reform has on an individual's log net-of-tax rate. Let the average ETI parameter for individuals in the reform cohort with initial income $X = x$, whose log net-of-tax rate changes by $\Phi = \phi$, be denoted by

$$\zeta(x, \phi) \equiv \mathbb{E}[\zeta \mid G = 1, X = x, \Phi = \phi].$$

for each ϕ and x.[5] We collect the group-specific average ETI parameters in $\bar{\zeta} = \{\zeta(x, \phi) : x \in \mathcal{X}, \phi \in \mathcal{P}\}$, where $\mathcal{X}$ and $\mathcal{P}$ is the support of X and Φ respectively. Our definition of a causal ETI estimand requires β to recover a positively weighted average of the group-specific average ETI parameters:

Definition 1 (Causal ETI estimand)**.** *The ETI estimand β is causal if, for any $\bar{\zeta}$:*

$$\beta = \sum_{x,\phi} \omega(x, \phi) \times \zeta(x, \phi)$$

with $\sum_{x,\phi} \omega(x, \phi) = 1$ and $\omega(x, \phi) \geq 0$ for all $x \in \mathcal{X}$ and $\phi \in \mathcal{P}$.

The requirement that the weights sum to 1 ensures that a causal ETI estimand recovers ζ when ζ is constant across individuals, while the non-negative weights are necessary to ensure that a causal ETI estimand is contained in the support of $\zeta(X, \Phi)$. For example, if the weights are negative for some values of x and ϕ, then β could be negative even if $\zeta(x, \phi)$ is positive for all x and ϕ.

Proposition 1 provides necessary and sufficient conditions for the ETI estimand to be causal. It shows that $h(Z)$ must only take two values and that the controls for initial income f must be sufficiently rich:

Proposition 1. *Suppose Assumptions 1 - 3 hold. Then, the ETI estimand β is causal if and only if $h(Z)$ takes exactly two values and*

$$\mathbb{E}\left[\mathbb{L}[Gh(Z) \mid G, X] \mid X\right] = \mathbb{E}[Gh(Z) \mid X] \tag{10}$$

where $\mathbb{L}[Gh(Z) \mid G, X]$ are the population fitted values from regressing $Gh(Z)$ onto $(G, m_1(X), ..., m_J(X))$.

While we refer to Appendix A for a formal proof, it is useful to observe the two distinct reasons why an ETI estimand may fail to be causal.

First, if $h(Z)$ takes more than two values, the ETI estimand will compare the earnings responses between individuals who experience (non-zero) changes in marginal tax rates of varying degrees. These comparisons could lead to negative weights if there is heterogeneity in the ETI parameters.[6] In contrast, if $h(Z)$ is binary, the ETI estimand only compares earnings changes between individuals treated by the reform ($Z > 0$) and untreated individuals ($Z = 0$), who do

[5] As a minor abuse of notation, we assume that $\zeta(x, \phi)$ is well-defined for all (x, ϕ), even if ϕ is not in the support of Φ given X, so that $\mathbb{P}(\Phi = \phi, X = x \mid G = 1) = 0$. This convention has no impact on our results.

[6] Callaway et al. (2025) formalizes the problem of variable treatment intensity in difference-in-differences, and shows that strong auxiliary assumptions are needed (e.g., constant effects) to give the estimate a causal interpretation.

not experience any change in marginal tax rates. These comparisons will not produce negative weights, even if the elasticities are heterogeneous.

Second, if f does not satisfy the richness condition in equation (10), it is insufficiently flexible to capture how counterfactual earnings and marginal tax rate growth vary with X given G. Thus, the excluded instrument $Gh(Z)$ in the first stage (4) may be correlated with the error term u^y in the outcome equation (3). As a result, the ETI estimand cannot be given a causal interpretation, even when ζ is homogeneous.

The richness condition in (10) requires f to be flexible enough to exactly reproduce the conditional mean of the instrument given initial income X. If f is saturated-in-X, this condition necessarily holds, since it includes a parameter for every value of X.[7] By contrast, it necessarily fails if f does not allow for discontinuous changes at $x = \bar{Y}$.

It is useful to observe that the condition in (10) is about how to model the conditional expectation function of the instrument given income controls, not "correctly" specifying the outcome equation or the first stage of the IV model. The proposition is therefore distinct from most of the literature on misspecification bias in instrumental variables estimation.

Proposition 1 considers the case where the ETI parameters ζ are allowed to vary across individuals. Corollary 1 specializes the result to the case of homogeneous ETI parameters. The necessary and sufficient condition is then that f satisfies the richness condition. A key implication is that the widely used polynomial specification of income controls cannot be given a causal interpretation, even if the ETI parameters are constant.

Corollary 1. *Suppose Assumptions 1 and 2 hold, and assume that ζ is homogeneous across individuals. Then, $\beta = \zeta$ if and only if f satisfies equation* (10).

2.5 *The* causal interpretation of the ETI estimand

It is important to emphasize that the criterion of an estimand to be causal in Definition 1 is a weak one. Thus, being causal may be necessary for the ETI estimand to be an interesting quantity, but it is not sufficient. For example, the definition does not preclude that all the weight is assigned to a group with a negative $\zeta(x, \phi)$, while the rest of the population receiving zero weight have positive ζ. The following corollary strengthens the result in Proposition 1 by showing *the* causal interpretation of the ETI estimand as a specific positively weighted average of group-specific ETI parameters $\zeta(x, \phi)$.

Corollary 2. *Suppose Assumptions 1 - 3 hold. Then, the causal ETI estimand equals*

$$\beta = \sum_{x,\phi} \overbrace{\omega(x, \phi)}^{\text{group-specific weights}} \times \underbrace{\zeta(x, \phi)}_{\text{group-specific average ETI}},$$

[7]Note that a specification of f that is saturated-in-X satisfies our richness condition in (10), even if there is no interaction between X and G. This is because we impose Assumption 1 (Common trends), which gives separability in how counterfactual earnings and marginal tax rate growth vary with X and G. Therefore, our condition (10) is weaker than the necessary and sufficient condition of rich covariates for a causal interpretation of IV estimands in Blandhol et al. (2022).

with weights given by

$$\omega(x,\phi) \equiv \frac{\mathbb{1}[x \geq \bar{Y}]\,\mathbb{P}(X=x)Var(G \mid X=x)\,\mathbb{P}(\Phi=\phi \mid G=1, X=x) \times \phi}{\sum_{x',\phi'} \mathbb{1}[x' \geq \bar{Y}]\,\mathbb{P}(X=x')Var(G \mid X=x')\,\mathbb{P}(\Phi=\phi' \mid G=1, X=x') \times \phi'},$$

that are positive and sum to one.

The corollary shows that the ETI estimand recovers a specific positively weighted average of group averages of ζ. The groups are mutually exclusive and defined by initial income X and the reform's effect on their marginal tax rate Φ. Conditional on initial income $X = x$, groups are weighted according to their size and how large changes in the marginal tax rate Φ they experience.

The ETI estimand aggregates the groups' average ETI parameters *across* initial income in proportion to how dispersed observations with $X = x$ are across cohorts G, as measured by $Var(G \mid X)$. These weights resemble how linear regression aggregates average treatment effects across covariates, and can be viewed as efficiency weights (see, e.g., Angrist, 1998).

2.6 Quantifying how the elasticities vary across the income distribution

Corollary 2 showed that the ETI estimand β recovers a positively weighted average of heterogeneous ETI parameters ζ across groups defined by initial income X and marginal tax rate change Φ. An important question for tax policy is how the ETI varies across the income distribution X. To analyze this question, we introduce the *local* ETI estimand $\beta(x)$, as defined by

$$\beta(x) \equiv \frac{\overbrace{\mathbb{E}[\Delta y \mid G=1, X=x]}^{\text{actual income growth}} - \overbrace{\left(\lambda_0^y + f^0(x;\lambda^y)\right)}^{\text{counterfactual income growth}}}{\underbrace{\mathbb{E}[\Delta\,\text{NTR} \mid G=1, X=x]}_{\text{actual tax rate growth}} - \underbrace{\left(\lambda_0^{\text{NTR}} + f^0(x;\lambda^{\text{NTR}})\right)}_{\text{counterfactual tax rate growth}}}, \tag{11}$$

for each $x \geq \bar{Y}$, where the function $f^0(\cdot,\lambda)$ is saturated-in-X, linear in the parameter vector λ, and λ_0^w and λ^w are given by

$$(\lambda_0^w, \lambda^w) \equiv \arg\min_{\lambda_0,\lambda} \mathbb{E}\left[\left(\Delta w - \lambda_0 G - f^0(X;\lambda)\right)^2 \mid GZ = 0\right], \tag{12}$$

for $w = y, \text{NTR}$.

Intuitively, the estimand exploits that Assumption 1 implies the earnings and marginal tax rate growth among untreated ($GZ = 0$) individuals can be used to recover the counterfactual growth of the treated ($GZ \neq 0$) individuals. Subtracting the counterfactual growth from the treated individual's actual growth yields their earnings response and the changes in their marginal tax rate due to the tax reform. Our next result shows that the local ETI estimand $\beta(x)$ is causal under the same assumptions that were necessary to give the ETI estimand β a causal interpretation:

Proposition 2. *Suppose Assumptions 1 - 3 hold. Then, the local ETI estimand $\beta(x)$ is causal and equals:*

$$\beta(x) = \sum_{\phi} \overbrace{\frac{\mathbb{P}(\Phi=\phi \mid G=1, X=x)\phi}{\sum_{\phi'} \mathbb{P}(\Phi=\phi' \mid G=1, X=x)\phi'}}^{\textit{weights reflecting group size and changes in marginal tax rates}} \times \underbrace{\zeta(x,\phi)}_{\textit{group-specific average ETI}}, \tag{13}$$

for each $x \geq \bar{Y}$. The weights are positive and sum to one.

Proposition 2 shows that the local ETI estimand $\beta(x)$ recovers a positively weighted average across the same group-specific averages of ζ as in Corollary 2.[8] However, only groups with initial income $X = x$ receive positive weights. Among the groups with $X = x$, the local ETI estimand weights the groups according to their size $\mathbb{P}(\Phi = \phi \mid G = 1, X = x)$ and how much their marginal tax rates are affected by the reform Φ.

3 Estimating elasticities of taxable income

We now apply the identification results from Section 2 to analyze a reform of the Norwegian tax system that reduced the marginal tax rates on middle and high incomes.

3.1 The Norwegian tax system and the 2006 tax reform

The Norwegian income tax system combines a flat tax on general income with a progressive surtax on labor earnings.[9] General income consists of both labor earnings and capital income, and is taxed at a flat rate of 28 percent after allowable deductions. Deductions – such as the wage-earner deduction, the personal allowance, and selected pension and interest deductions – apply exclusively to the general income tax base. The surtax on labor earnings consists of a 7.8 percent social security tax and several income brackets with increasing marginal tax rates. The overall marginal tax rate on labor earnings thus combines the flat general income tax and the progressive surtax. Throughout the paper, we refer to the corresponding tax base as labor earnings or taxable income.

We study a tax reform that significantly reduced the marginal tax rate on middle- and high-income earners. The reform was partially introduced in 2005 and took full effect on January 1, 2006. This reform has been previously analyzed by Vattø (2020), who estimates an average ETI of 0.15 (see also Thoresen and Vattø, 2015 and Berg and Thoresen, 2020).

Figure 2 plots the marginal tax rates by labor earnings before (2003) and after (2006) the tax reform was fully implemented. The graph shows that the reform reduced marginal tax rates on medium and especially high earnings: From 2003 to 2006, the marginal tax rate in the (second) highest bracket was reduced from 55.3 to 47.8 (49.3 to 44.8) percent. The sizable rate cuts meant that most workers in the upper half of the income distribution experienced large reductions in their marginal tax rates.

A potential concern with using the 2006 reform as an instrument for marginal tax rates is that a dividend tax was introduced at the same time. The introduction of the dividend tax may have reduced the incentives for self-employed individuals and owners of closely held corporations to report labor earnings as capital income. As a consequence, observed changes in labor earnings could reflect income shifting in response to the dividend tax rather than earnings responses due to changes in the marginal tax rates. We examine this possibility in two ways. First, we exclude self-employed individuals and owners of closely held corporations from the sample. Second, we

[8] The expression is similar to equation (23) in Mogstad and Torgovitsky (2024) with the additional restriction that the potential earnings functions are linear. See their discussion of how it relates to the *average causal response* from Angrist and Imbens (1995).

[9] See Mogstad et al. (2025) for a detailed discussion of the Norwegian tax system and labor market.

estimate the ETI using total market income (the sum of labor and capital income) as the outcome variable. If the observed responses were driven entirely by income shifting, the reform should have no effect on total market income. It is reassuring to find that the estimates do not materially change in these specification checks.

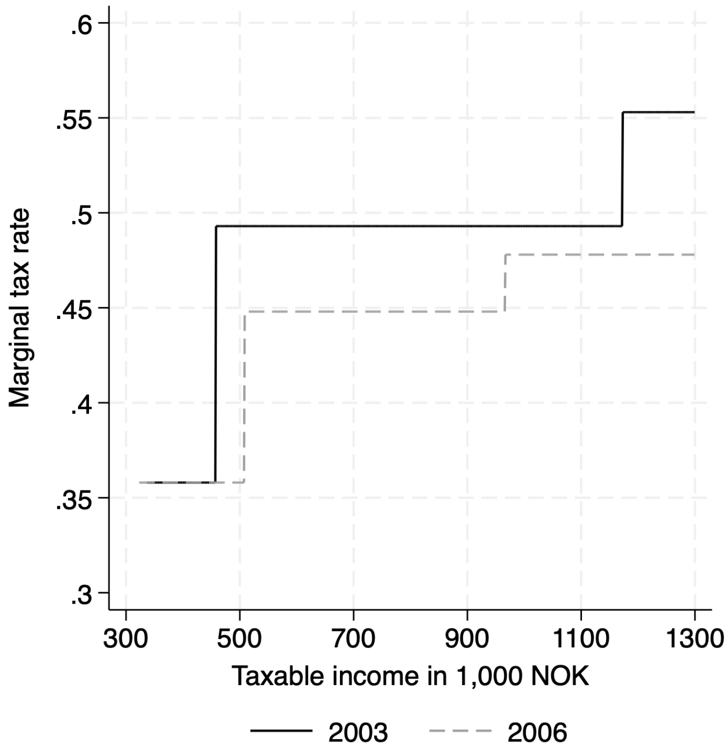


Figure 2: Marginal tax rates in 2003 and 2006

Notes: This figure shows the marginal tax rates that apply to different income tax brackets in 2003 and 2006. The tax bracket thresholds are measured in 2018 NOK.

3.2 Data and sample

Our empirical analysis is based on several administrative data sources, which we link using unique identifiers for individuals and households. Marginal tax rates are constructed using the same tax simulation model as Statistics Norway and the Norwegian Ministry of Finance. This results in a panel dataset covering the full Norwegian population in the period 1998–2017. The dataset includes detailed information from income tax returns as well as individual characteristics such as age, sex, educational attainment, and marital status.

To explain how our estimation sample is constructed, it is useful to recall the discussion of the ETI estimands in Section 2. There, we showed that the ETI estimand compares how earnings and marginal tax rates evolve over periods when the tax system changed (2003 vs 2006) with how they evolve over periods when the tax system did not change (1999 vs 2002). For each individual in the reform cohorts, we constructed one earnings difference by subtracting earnings in 2003 from earnings in 2006, and similarly for the marginal tax rates. Correspondingly, for each individual in the placebo cohorts, we constructed one earnings difference by subtracting earnings in 1999 from earnings in 2002.

We can increase the sample size by considering differences between years other than 2006 and 2003 (reform cohorts) and 2002 and 1999 (placebo cohorts). The sample size of the reform cohorts can be increased by including the differences between 2002 and 2005 and between 2004 and 2007. The sample size of the placebo cohorts can be increased by including the differences between 2000 and 2003 and between 2001 and 2004. Our empirical analysis includes all six three-year differences. We refer to the differences over periods where the tax system changed (2002 vs 2005, 2003 vs 2006, and 2004 vs 2007) as the reform differences. The differences over periods with no changes in the tax system (1999 vs 2002, 2000 vs 2003, and 2001 vs 2004) are referred to as the placebo differences. Our point estimates do not materially change if we instead focus on particular reform and placebo differences.

Our estimation sample is obtained by imposing four restrictions. First, we only include individuals aged 25 to 61 whose wage earnings are their primary source of labor income. Students and individuals receiving pensions or unemployment benefits are excluded. Second, we exclude differences with initial income below 300,000 (15th percentile). Excluding individuals with low incomes is common in the literature since mean reversion tends to be most pronounced at the bottom of the income distribution (see Saez et al., 2012). Third, we exclude differences with initial income above 1,100,000 (97th percentile). This restriction ensures that placebo differences are not affected by a top-income tax bracket introduced in 2000 for very high incomes (above 1,100,000). Fourth, we exclude placebo differences with initial income between 400,000 and 475,000. This is done to minimize the extent to which placebo differences are affected by small year-to-year changes in tax bracket thresholds.

Our resulting estimation sample contains about 4 million observations of Δy and 1.1 million unique individual wage earners. Summary statistics for our baseline and estimation samples are provided in Table A1 in Appendix B.

3.3 Estimates of average ETIs

To implement the ETI estimands, we use the data on the reform and placebo differences to estimate the model in equations (3) and (4) with two-stage least squares. Our specification includes indicator variables for each of the six three-year differences, and interacts the simulated instrument $h(Z)$ with a dummy that equals 1 if the observation is a reform difference and zero if it is a placebo difference. Table 1 reports the resulting estimates under different choices of h and f.

The first column of Panel A reports estimates from our preferred specification, which puts $h(Z) = \mathbb{1}[Z > 0]$ and controls for initial income using a full set of percentile dummies in X. The first stage estimates reported in the first row are precise and show that the reform increases the marginal net-of-tax rate by around 5 percent. The F-statistic on the excluded instrument is large, indicating that the first stage is very strong. The reduced form, reported in the second row, is also precisely estimated and shows that the tax reform increases earnings by 1.1 percent.

The ETI estimate, reported in the third row, is 0.21 (with a standard error of 0.012). The specification that produces this estimate controls for initial income non-parametrically and makes no comparisons of earnings responses between individuals who experience (non-zero) changes in marginal tax rates of varying degrees. Thus, it can be interpreted as a positively weighted average of ETI parameters under the assumptions in Proposition 1. If the ETI parameters are constant across individuals, the estimate implies that a ten percent increase in the marginal net-of-tax rate raises earnings by 2.1 percent.

Table 1: Main results

Panel A. Specification of income controls f: dummies for X in bins.				
Instrument		Binary: $h(Z) = \mathbb{1}[Z > 0]$		Continuous: $h(Z) = Z$
Income bins		Percentiles	Above/below $\bar{Y}$	Percentiles
First stage	Coef.	0.0504	0.0493	0.3063
	SE	0.0002	0.0002	0.0017
Reduced form	Coef.	0.0105	0.0092	0.0492
	SE	0.0006	0.0005	0.0039
ETI	Coef.	0.2091	0.1858	0.1605
	SE	0.0124	0.0109	0.0130
Panel B. Specification of income controls f: linear splines.				
Instrument		Binary: $h(Z) = \mathbb{1}[Z > 0]$		Continuous: $h(Z) = Z$
Spline knots		Deciles	Above/below $\bar{Y}$	Deciles
First stage	Coef.	0.0505	0.0647	0.3012
	SE	0.0002	0.0002	0.0016
Reduced form	Coef.	0.0105	0.0099	0.0478
	SE	0.0006	0.0005	0.0038
ETI	Coef.	0.2077	0.1531	0.1587
	SE	0.0122	0.0076	0.0129

Notes: The table contains two-stage least squares estimates of β from the model in (3) and (4) under different choices of h and f. All regressions include time dummies for each year and a dummy that equals 1 if the observation is a reform difference and zero if it is a placebo difference.

The second and third columns of Panel A report estimates of specifications that fail to satisfy the conditions for a causal interpretation of the ETI estimand. The second column corresponds to the triple difference estimand in (2). It correctly specifies the instrument, but does not control flexibly for initial income, resulting in an estimate that understates the average ETI parameter.

The third column reports estimates from a commonly used ETI estimand that is causal only under the assumption of constant ETI parameters. It controls flexibly for initial income but specifies the instrument as $h(Z) = Z$. This means that it compares earnings responses between individuals who experience (non-zero) changes in marginal tax rates of varying degrees. The resulting estimate is equal to 0.16 and significantly understates the average ETI parameter. Thus, we can reject the assumption of homogeneous ETI parameters.

Panel B of Table 1 reports estimates of specifications that control for initial income through splines instead of dummies for income bins. We report these estimates because spline specifications are frequently used in the ETI literature. We find that whether one flexibly controls for income through splines or income bins does not materially affect the results. We conclude that when the instrument is binary, and the income controls f are reasonably flexible, the ETI estimates are robust to the exact specification of f.

3.4 Placebo checks and robustness

One possible concern with the validity of the estimates in Table 1 is that Assumption 1 (Common trends) may not hold. To address this concern, we re-estimate the model in equations (3) and (4) on the placebo differences, treating one of the placebo differences as if it were a reform difference. We perform this procedure for all three placebo differences and report the averages in Table 2.

The resulting placebo reduced form, reported in the first column, is indistinguishable from zero. To obtain a placebo ETI estimate, one would need to divide the placebo reduced form by a first stage. However, because the placebo differences are not affected by any tax reform there is no first stage. We instead divide the placebo reduced form by the first stage estimate from the corresponding specification in Table 1. The resulting placebo ETI estimate is multiple orders of magnitude smaller than the actual ETI estimate.

Table 2: Placebo responses

Instrument		Binary: $h(Z) = \mathbb{1}[Z > 0]$	
Income bins		Percentiles	Above/below $\bar{Y}$
Placebo reduced form	Coef.	0.00006	0.00002
	SE	0.00001	0.00001
Placebo ETI	Coef.	0.00110	0.00045
	SE	0.00011	0.00009

Notes: This table presents placebo estimates of the reduced form from the model in equations (3) and (4). The placebo ETI is the reduced form divided by the actual first stage estimate from the corresponding specification in Table 1.

Another concern with the estimates in Table 1 is that the tax reform could induce extensive margin responses, potentially leading to compositional differences between the treatment and control groups. In our setting, this is unlikely to be a serious concern. First, the tax reform affects individuals with relatively high incomes, making extensive-margin responses unlikely. Second, our estimation sample consists of individuals employed in the years preceding the reform. Since the reform decreases participation tax rates, we do not expect these individuals to stop working as a result of the reform.

To assess the prevalence of employment responses to the reform, we estimate the reduced form of the regression model in equations (3) and (4) on a sample that includes the unemployed, where we have replaced the outcome with a binary variable equal to one if the individual works and zero otherwise. Consistent with our discussion above, Table A2 in Appendix B shows that the reform increases employment by approximately 0.06 percentage points, indicating a small but non-zero effect on employment. To understand how these small employment effects could bias the ETI estimates, we develop bounds on the ETI while allowing for employment responses to the reform in Appendix C. The estimated bounds are reported in Appendix Table A2. The bounds are centered around our point estimate from above, with an upper bound equal to 0.28 and a lower bound equal to 0.18. The relatively narrow bounds reflect the small employment response. We conclude that employment responses are not a key confounding factor for our ETI estimates.

To address the potential concern related to the introduction of a dividend tax as part of the tax reform (see discussion in Section 3.1), we now consider how excluding self-employed individuals and owners of closely held corporations affects our estimates. Column (i) of Table A3 in Appendix

B shows an estimated ETI of 0.208. Column (ii) reports estimates using total market income (the sum of labor and capital income) as the outcome. The resulting ETI estimate of 0.18 is close to the baseline estimate of 0.21, indicating that our result is not driven by income shifting.

Our baseline regression results put the same weight on each individual. However, if the parameter of interest is the revenue effects of tax policy, one can argue that the regression needs to weight individuals by their initial income (see Gruber and Saez (2002)). Column (iii) in Table A3 shows that weighting by X leads to a slightly larger estimated ETI of 0.23.

3.5 Estimates of ETIs across the income distribution

We now turn to assessing how the estimated ETI parameters vary across the distribution of initial income. We implement the estimator for $\beta(x)$ in two steps. The first step estimates the counterfactual income and tax rate growth in equation (12) using 100 percentile bins in X, while the second step estimates the numerator and denominator of equation (11) separately using a local regression.

Panel (a) of Figure 3 plots the resulting estimates. The estimated ETI is less than 0.1 for incomes around 450,000 (median), increases steadily to around 0.35 for incomes close to 750,000 (90th percentile), and exceeds 0.5 for incomes around 900,000 (95th percentile). The variation in the estimates of $\beta(x)$ across x is significant, both statistically and economically, and means that the (weighted-) average elasticity recovered by the causal ETI estimand β is far from sufficient to assess how changes in marginal tax rates affect earnings and tax revenue.

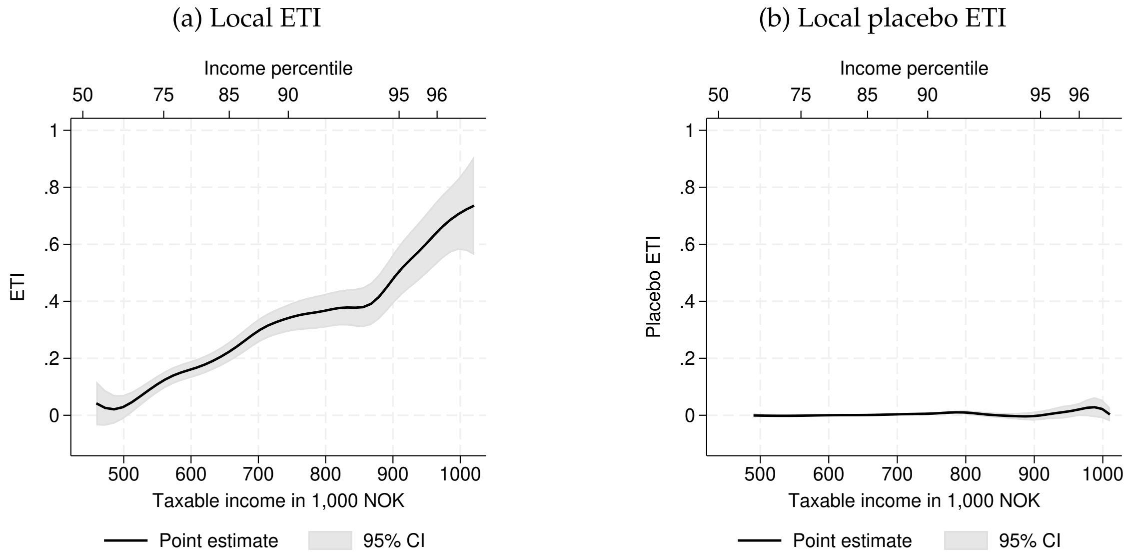


Figure 3: Local ETI estimates.

Notes: The figure plots actual and placebo estimates of the local ETI across the distribution of initial income. Panel (a) plots the actual local ETI obtained by first estimating the counterfactual income and tax rate growth using 100 percentile bins of X, then estimating the numerator and denominator of equation (11) separately using a local regression. To obtain the local placebo ETIs in Panel (b), we re-estimate the reduced form in the numerator of equation (11), treating one of the placebo differences as if it were a reform difference. We perform this procedure for all three placebo differences and take the average. The local placebo ETI estimates are then obtained by dividing these placebo-reduced forms by the actual first stage. 95 percent confidence intervals are shown, with standard errors obtained by bootstrapping the entire estimation procedure using 500 replications.

Panel (b) of Figure 3 assesses the common trend assumption behind the local ETI estimand by estimating a placebo version of the local ETI estimand following the same steps as for the placebo estimates reported in Table 2. The placebo estimates remain close to zero throughout the income distribution.

To assess how employment responses could affect our local ETI estimates, Panel (a) of Figure A1 in Appendix B plots the local employment ETIs. Employment responses are small but non-zero across most of the income distribution. To assess how they could influence our estimates, Appendix C derives bounds on the local ETIs when allowing for employment responses. As shown in Panel (b), our point estimates lie within the estimated bounds, which are relatively narrow, reflecting the small employment responses. We conclude that employment responses are not a key confounding factor for our local ETI estimates.

4 Using ETIs to learn about labor supply elasticities

We now use the ETI estimates from above to learn about labor supply elasticities.

4.1 A labor supply model

We consider a labor supply model where workers have convex preferences over K consumption goods C_k and L margins of labor supply Y_l. The L margins of labor supply produce pre-tax income Y according to $Y = F(Y_1, ..., Y_L)$, which is concave and strictly increasing in each of its arguments.

Multiple dimensions of labor supply allow for the possibility that individuals can affect their earnings through many different choices, including hours of work, effort on the job, and firm and occupation choices. Introducing multiple consumption goods is useful for computing total marginal tax rates, as it accommodates differential taxation, such as varying value-added tax rates, across goods.

If the income tax system were linear with marginal tax rate τ and transfer R, the individual's utility maximization problem would be,

$$\max_{C_1,...,C_K,Y_1,...,Y_L} U(C_1, ..., C_K, Y_1, ..., Y_L) \qquad \text{subject to } \sum_{k=1}^{K}(1+\tau_k)P_k C_k \leq I, \\ I = (1-\tau)Y + R + B, \text{ and } Y = F(Y_1, ..., Y_L), \tag{14}$$

where B is unearned income, I is consumption expenditure, and P_k and τ_k are the price of and tax on consumption good k, respectively.

As noted by Feldstein (1999), the revenue- and efficiency effects of taxation typically depend on how tax policy affects earnings, but not on the specific margins through which individuals adjust their behavior. We therefore focus on the earnings choice, which – suppressing its dependence on prices and consumption tax rates – can be written as,

$$Y^u(\tau, R+B) \equiv F(Y_1^u(\tau, R+B), ..., Y_L^u(\tau, R+B)), \tag{15}$$

where $Y_l^u(\tau, R+B)$ is the optimally chosen l-th labor supply component. This labor supply function allows for defining the standard labor supply elasticities,

$$\varepsilon^u \equiv \frac{\partial Y^u}{\partial 1-\tau}\frac{1-\tau}{Y^u}, \qquad \eta \equiv (1-\tau)\frac{\partial Y^u}{\partial(R+B)} \in [-1,0], \qquad \varepsilon^c \equiv \varepsilon^u - \eta \geq 0, \tag{16}$$

where ε^u and ε^c denote the uncompensated and compensated earnings elasticity, η denotes the income effect, and the relationship between (un)compensated elasticities and the income effect is given by the Slutsky equation. The restriction that $\eta \in [-1, 0]$ follows from the Engel aggregation condition by assuming that consumption and leisure are normal goods.

In reality, the Norwegian tax system is piecewise linear. To accommodate this feature, we follow the argument of Hall (1973): even if the actual tax system T_d is non-linear, convex preferences ensure that individuals behave as if they were facing the following linear budget constraint,

$$I = (1-\tau(d))Y + R(d) + B \text{ for } d = 0, 1, \tag{17}$$

with $R(d) \equiv \tau(d)Y(d) - T_d(Y(d))$, and $Y(d)$ and $\tau(d)$ defined as in Section 2.3. This argument implies that the potential earnings function $Y(\tau(d), d)$ in equation (7) can be viewed as the solution to the worker's labor supply problem, subject to the linear budget constraint defined in equation (17):

$$Y(\tau(d), d) \;=\; Y^u(\tau(d), R(d) + B) \text{ for } d = 0, 1. \tag{18}$$

4.2 Recovering labor supply elasticities

The previous subsection forged a tight link between the potential earnings function and the model of labor supply. Having established this link, we now use the ETI estimates (and other specific features of the data) to draw inferences about labor supply elasticities.

We first consider the case with no income effects, a restriction typically imposed in the models used to interpret ETI estimands. Our next result shows that this model restriction ensures that the exclusion restriction in Assumption 2 is satisfied and, therefore, lets us interpret the individual ETI parameters ζ as compensated earnings elasticities ε^c:

Proposition 3. *Assumption 2 is satisfied if and only if there are no income effects. If there are no income effects, then $\zeta = \varepsilon^c$.*

An immediate implication of the result is that the ETI estimands β and $\beta(x)$ in Section 2 recover positively weighted averages of compensated earnings elasticities, provided the instrument is binary and the specification of the income controls f is unrestricted.

Allowing for income effects violates the exclusion restriction in Assumption 2, since the tax reform can then affect labor supply through changes in both the marginal and the average tax rates. The following proposition clarifies the identification problem that arises due to income effects. It expresses ε^c and ε^u in terms of the local ETI estimands $\beta(x)$ and a bias term that would be observable in data if the income effects were known or could be estimated:

Proposition 4. *Suppose that ε^c and η are homogeneous conditional on X, and let $\varepsilon^c(x)$ and $\eta(x)$ denote the compensated elasticity and income effect for individuals with $X = x$. If Assumption 1 holds and there* *is no bracket switching, then, up to a first-order approximation,*

$$\varepsilon^c(x) = \beta(x) + B(x)\eta(x), \tag{19}$$

$$\varepsilon^u(x) = \beta(x) + (B(x) + 1)\eta(x), \tag{20}$$

where $B(x)$ is estimable and equal to

$$B(x) \equiv \frac{\mathbb{E}\left[\frac{T_1(Y_t)-T_0(Y_t)}{(1-T_1'(Y_t))Y_t} \mid G = 1, X = x\right]}{\mathbb{E}[\Delta\,\mathrm{NTR} \mid G = 1, X = x] - (\lambda_0^{\mathrm{NTR}} + f^0(x; \lambda^{\mathrm{NTR}}))}, \tag{21}$$

and $x \geq \bar{Y}$.

A key insight from Proposition 4 is that the local ETI estimand neither recovers the compensated elasticity nor the uncompensated elasticity if one allows for income effects, even if one assumes no bracket switching and that elasticities are homogeneous conditional on initial income X.[10] However, it also shows that the bias term is a multiplicatively separable function of the income effects $\eta(x)$ and an observable term $B(x)$. Therefore, if $\eta(x)$ were known or could be estimated, one could recover $\varepsilon^c(x)$ and $\varepsilon^u(x)$ from observed data. This observation motivates and guides the analysis in the remainder of the paper, where we consider different approaches to point identify or bound the compensated and uncompensated earnings elasticities.

4.3 No additional assumption bounds on labor supply elasticities

In order to use Proposition 4 to construct bounds on labor supply elasticities, it is useful to recall that the Engel aggregation condition implies that $\eta(x)$ is theoretically bounded between -1 and 0. Thus, evaluating equations (19) and (20) for $\eta(x) \in \{-1, 0\}$ produces bounds on compensated and uncompensated earnings elasticities without imposing assumptions other than those stated in Proposition 4.

The light grey areas of Figure 4 plot these bounds using the estimates of $\beta(x)$ reported in Figure 3 and estimates of $B(x)$ from the cross-sectional earnings distribution. Panel (a) provides two insights about compensated elasticities. First, the compensated elasticities for individuals with incomes equal to 650,000 (85th percentile), 750,000 (90th percentile), and 950,000 (96th percentile) are at least $0.2, 0.35$, and 0.7, respectively. Second, the bounds rule out larger than modest compensated elasticities for middle-income individuals: the upper bound for individuals with income around 450,000 (median) equals 0.5. Taken together, these insights imply that compensated elasticities increase by at least 40 percent as income increases from the median to the 96th percentile.

Panel (b) plots the corresponding bounds for the uncompensated elasticities. A key result is that the bounds are highly informative for high-income individuals, implying a relatively large uncompensated elasticity between 0.3 and 0.5 for individuals with incomes around 900,000 (95th percentile). A second result is that we can rule out that income effects dominate substitution effects for all individuals with income above 750,000 (90th percentile). Finally, the bounds rule

[10] The result in Proposition 4 does not invoke Assumption 3. It is no longer needed since it is assumed that elasticities are homogeneous among individuals with the same initial income X. For tractability, Proposition 4 also imposes no bracket switching, which will necessarily hold if the reform under consideration is only changing the top income tax rate.

out uncompensated elasticities above 0.1 for individuals with incomes below 550,000 (the 75th percentile).

(a) Compensated elasticity (b) Uncompensated elasticity

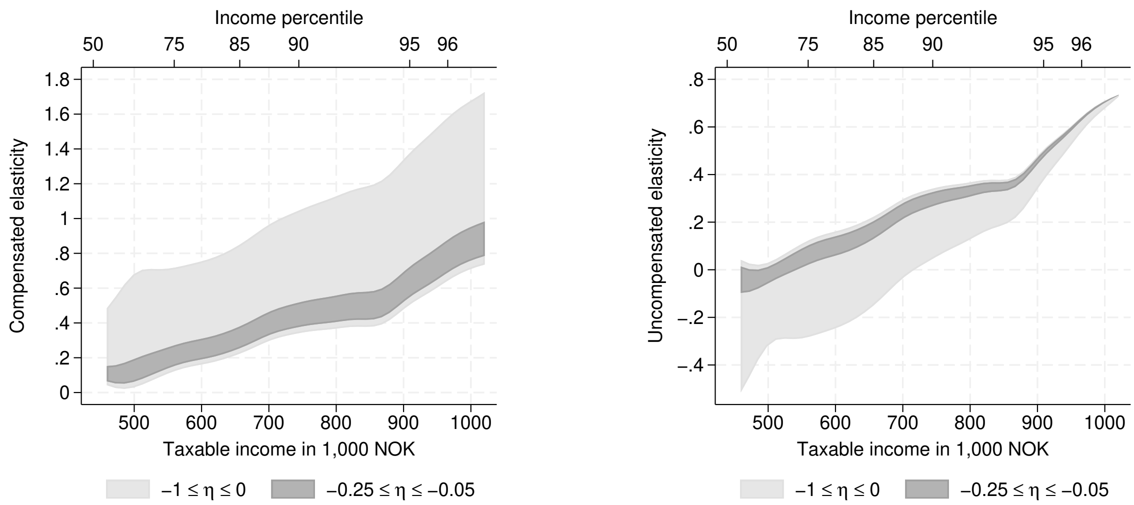


Figure 4: Average (un)compensated earnings elasticities across initial income X.

Notes: The figure plots bounds on the compensated and uncompensated labor supply elasticities over the distribution of initial income. The light grey area in Panel (a) plots the bounds on $\varepsilon^c(x)$ using that $\eta \in [-1, 0]$, while the darker grey area imposes that $\eta \in [-0.25, -0.05]$. The light grey area in Panel (b) plots the bounds on $\varepsilon^u(x)$ using that $\eta \in [-1, 0]$, while the darker grey area imposes that $\eta \in [-0.25, -0.05]$.

The bounds in light grey in Figure 4 can be tightened considerably by ruling out implausibly small and large income effects. The dark grey areas in Figure 4 represent the bounds implied by assuming that $\eta \in [-0.25, -0.05]$. They show that this additional assumption tightens both the upper bound on the compensated elasticities and the lower bound on the uncompensated elasticities considerably. The implied bounds suggest an (un)compensated elasticity between 0.05 and 0.15 (between -0.1 and 0) for incomes around 450,000. The (un)compensated elasticity is bounded between 0.4 and 0.5 (0.25 and 0.35) for incomes around 750,000.

Although the bounds in Figure 4 are highly informative, the (un)compensated elasticities remain partially identified. We next consider two ways to move from partial to point identification. One possibility, which we consider in subsection 4.4, is to assume constant income and substitution effects across the income distribution. Another possibility, which we consider in Section 5, is to combine the ETI estimands with external estimates of income effects.

4.4 Point identification with constant income and substitution effects

As shown by Gruber and Saez (2002), it is possible to use variation from tax reforms to jointly estimate income and substitution effects under the assumption that the compensated elasticity ε^c and the income effect η are constant across all individuals. Assuming homogeneity, Proposition 4 implies that the local ETI estimands $\beta(x)$ relate to the constant ε^c and η through,

$$\beta(x) = \varepsilon^c - \eta B(x). \tag{22}$$

A natural way to estimate (ε^c, η) is the following least squares estimator

$$(\hat{\varepsilon}^c, \hat{\eta}) = \underset{\varepsilon^c,\eta}{\arg\min} \sum_{x \geq \bar{Y}} \mathbb{P}(X = x \mid G = 1) \left(\hat{\beta}(x) - \varepsilon^c + \eta \hat{B}(x)\right)^2. \quad (23)$$

An estimator for the uncompensated elasticity is then $\hat{\varepsilon}^u = \hat{\varepsilon}^c + \hat{\eta}$.

Since equation (22) must hold for any x, ε^c and η are (over)identified if $\beta(x)$ and $B(x)$ are observed for (more than) two values of x. Even with only two values of x, the homogeneity assumption can be tested by examining if the labor supply parameters satisfy the theoretical restrictions $\varepsilon^c > 0$ and $\eta \in [-1, 0]$. In the overidentified case, the sharp test would not only use these theoretical restrictions, but also that the sum of squared residuals in (23) is zero.

Using the same estimates of $\beta(x)$ and $B(x)$ as in Figure 4, we solve the problem in (23) and obtain $\hat{\varepsilon}^c = -0.7274$ and $\hat{\eta} = -1.4131$. Since the estimated compensated elasticity is negative and the income effect is below -1, the homogeneity assumption is clearly at odds with the data. In fact, it can be statistically rejected: the bootstrapped standard error on ε^c is 0.081, meaning that the null of $\varepsilon^c \geq 0$ is rejected at the one percent level.

To understand why the assumption is rejected, it is useful to recall that the tax reform we consider decreased the top income tax rate. As a result, high-income individuals experienced a larger reduction in average tax rates than middle-income individuals. The assumption of constant ε^c and η therefore implies that earnings responses should decrease across the income distribution, in sharp contrast with Figure 3. Thus, we conclude that the assumption of constant income and substitution effects poorly approximates individual labor supply behavior, at least in our context.

5 Combining ETIs with external estimates of income effects to learn about labor supply elasticities

Proposition 4 shows that the compensated $\varepsilon^c(x)$ and uncompensated elasticities $\varepsilon^u(x)$ can be point identified from the local ETI estimands $\beta(x)$ if income effects $\eta(x)$ are known or can be estimated externally for each income level x. Motivated by this result, we now use Norwegian data on lottery winnings to measure income effects. We begin this analysis by estimating earnings and employment responses to lottery-induced changes in unearned income. Next, we show how these responses allow us to infer the the intensive margin earnings responses to unearned income that we need to point identify the (un)compensated elasticities.

5.1 Data and sample

Our empirical analysis combines multiple administrative data sources linked through unique personal and household identifiers. We supplement the tax records with data on unearned income and lottery winnings. The resulting panel covers the entire Norwegian population from 1995 to 2018. We refer to Eika et al. (2020) for a detailed description of the data.

Our sample selection mirrors the construction of the estimation sample for the ETI: We restrict our sample to winners aged 25–61 in the year before the win, require wage earnings to be their primary labor income source, and exclude students and individuals receiving pensions or unemployment benefits. Since individuals are only required to report winnings of NOK 100,000 or more, we restrict our sample to those who won at least this amount.[11] The final sample includes more than 18,000 unique winners across 23 years. The median prize is 347,000 NOK.

5.2 Measuring unearned income

To define unearned income and describe how it can be measured in the data, it is useful to state the households' intertemporal budget constraint. Letting A_{kt} and P^{A}_{kt} denote the holding and price of asset k at the beginning of period t, it can be written as

$$C_t + \sum_k P^{\mathrm{A}}_{kt} A_{kt+1} = Y_t - T(Y_t) + Y^{\mathrm{K}}_t + \sum_k P^{\mathrm{A}}_{kt} A_{kt},$$

where C_t is period t consumption expenditure and Y^{K}_t is after-tax capital income. By rearranging, we can express unearned income B_t as

$$B_t \equiv C_t - \Big(Y_t - T(Y_t)\Big) = Y^{\mathrm{K}}_t - \sum_k P^{\mathrm{A}}_{kt}\Big(A_{kt+1} - A_{kt}\Big), \tag{24}$$

where the last term is what Eika et al. (2020) refers to as net savings.[12] As these authors point out, net savings can be measured directly from data on sales and purchases of assets without having to measure or make assumptions about capital gains or changes in net wealth. Equation (24) thus implies that unearned income can be measured by observing capital incomes and asset transactions.

Following Eika et al. (2020), we use our detailed data to construct household-level measures of capital income and net savings. These measures allow us to compute unearned income directly using equation (24), which we then convert to a per-adult measure for consistent comparison between single and married households. This means that our rich data allow us to *observe* how winners allocate their wealth over time, eliminating the need to rely on the annuitization or capitalization approaches used in Imbens et al. (2001) and Golosov et al. (2024).

5.3 Research design and estimation

Before describing our research design, it is useful to introduce some terminology. We refer to all individuals who won a lottery in a given calendar year g as a cohort, and denote an individual's cohort by G. The event time t for cohort g corresponds to calendar year $g + t$, where t can be positive or negative depending on whether we look at outcomes before or after winning a lottery. We use the year before the lottery win as the reference year in the event study and refer to this year as the pre-win year. Subscripts now denote calendar year.

We explain our research design by showing how we recover the earnings effect of winning a lottery prize. We employ a difference-in-differences design that compares the evolution of earnings over time for individuals who have already won the lottery with those who have yet to win. For each event time t, we compare the earnings changes between years $g - 1$ and $g + t$ for cohort g to changes over the same two years for cohorts that will receive the prize at a later date, so that

[11] This threshold has been in place since 2007. Before 2007, the threshold was NOK 10,000.

[12] This definition of unearned income is consistent both with intertemporal two-stage budgeting in the absence of liquidity constraints and with the presence of liquidity constraints (Arellano and Meghir, 1992; Blundell and Macurdy, 1999; Blundell and Walker, 1986 and MaCurdy, 1983).

the comparison group remains untreated at both points in time. Formally, we define the reduced form difference-in-differences estimand by:

$$\text{RF}_{g,t}(x) \equiv \underbrace{\mathbb{E}\left[Y_{g+t} - Y_{g-1} \middle| G = g, Y_{g-1} = x\right]}_{\text{earnings change for year } g \text{ winners}} - \overbrace{\mathbb{E}\left[Y_{g+t} - Y_{g-1} \middle| G > g + t, Y_{g-1} = x\right]}^{\text{earnings change for later-than } g+t\text{-winners}}. \quad (25)$$

This estimand captures how earnings change for lottery winners relative to individuals with the same initial income Y_{g-1} who have not yet won. We show in Appendix D that $\text{RF}_{g,t}(x)$ with $t \geq 0$ recovers the earnings effect of winning the lottery under a parallel trends assumption.

In estimation, we implement the conditioning on Y_{g-1} using kernel weights and estimate $\text{RF}_{g,t}(x)$ separately for each cohort g and event time t. The estimands $\text{RF}_{g,t}(x)$ are then aggregated across cohorts using cohort-size weights. When aggregating across event time, we weight each event time equally. For each x, we estimate all cohort-by-event-time parameters in a single, fully interacted specification, which allows us to compute standard errors for aggregated effects using the delta method. All specifications include flexible controls for age to account for any systematic age differences between earlier and later winners.

5.4 The effect of winning the lottery on earnings and employment

Panel (a) of Figure 5 presents unconditional estimates of the effect of winning on winners' earnings for different event times. The estimates show no evidence of systematically different pre-trends between current and future winners. They reveal that earnings decline sharply in the first two years after winning and then stabilize at a lower level: After 3 years, average earnings are reduced by around 18,000.

The average effects in Panel (a) could mask considerable heterogeneity. Panel (b) shows how pre- and post-win estimates vary across the distribution of initial income and shows that pre-trends are similar between treatment and controls throughout the distribution. Following the win, we observe a substantial earnings reduction of around 10,000 – 20,000 that increases in magnitude with initial income.

Figure A2 in Appendix B plots the corresponding figures for the employment response. Panel (a) shows no evidence of systematic differences in employment time trends between current and later winners. After winning, the employment of winners declines over time and is reduced by almost 1.5 percentage points after 5 years. Panel (b) plots the pre- and post-event employment response estimates across the distribution of initial income. The employment responses are more pronounced at lower levels of initial income. While winning reduces employment by 0.8 percentage points among winners with initial income in the 10th percentile, the corresponding response among winners in the 90th percentile is roughly half as large.

5.5 The earnings response to unearned income

The magnitudes of the effects estimated in the previous section may be hard to interpret, as the observed responses to windfall gains could vary across individuals depending on a number of factors, such as the age at which the individual wins and her savings behavior. Motivated by this, we now turn to estimating how earnings respond to plausibly exogenous variation in unearned

(a) Unconditional effect on earnings

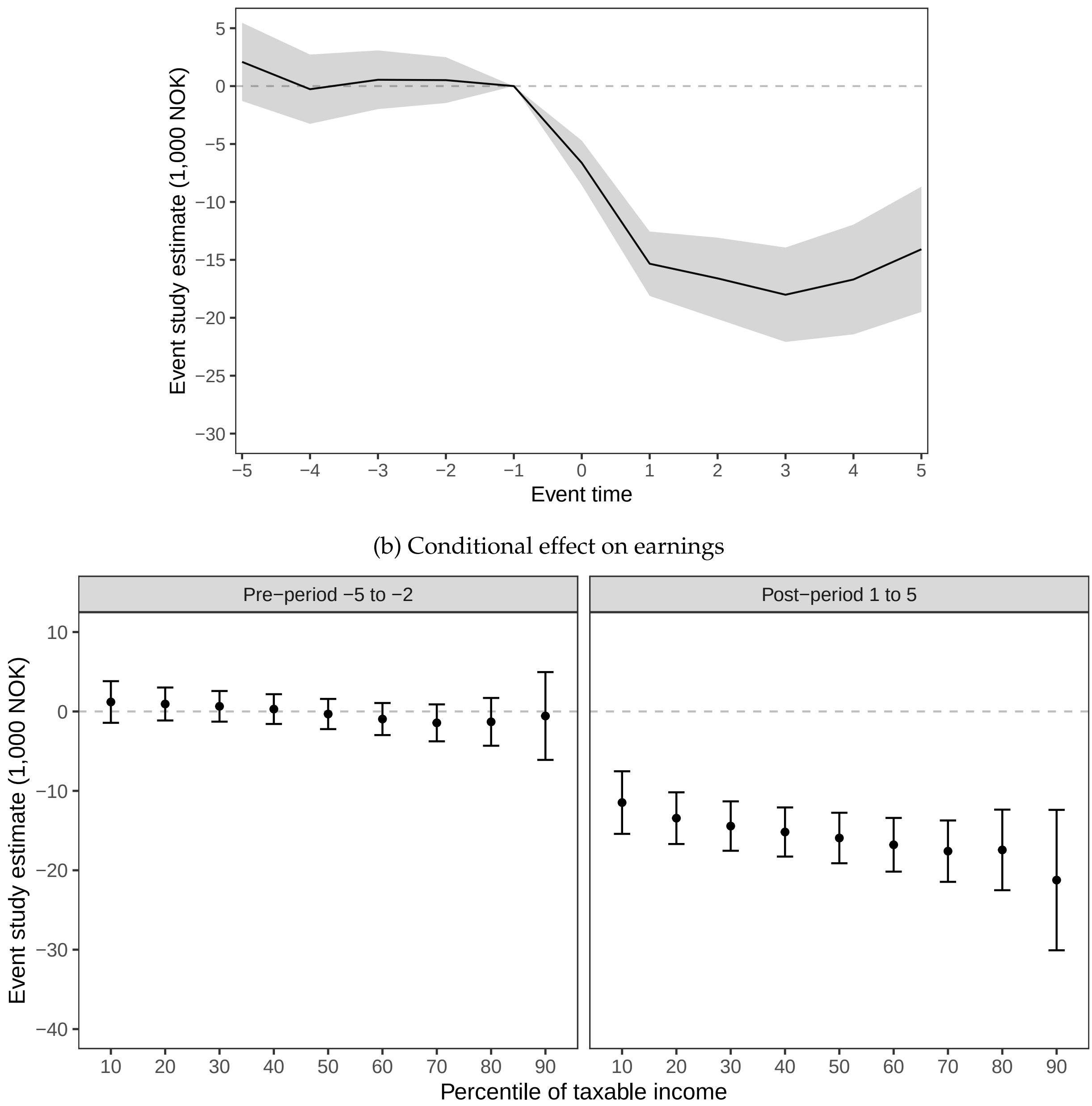


Figure 5: Earnings effects of winning

Notes: This figure presents estimates of the effect of winning the lottery on winners' earnings. Panel (a) presents unconditional estimates of the effect of winning the lottery on winners' earnings for each event time t. The estimates correspond to the sample analogue of equation (25), evaluated without conditioning on X, controlling for age, and averaged across cohorts using cohort-size weights. The estimates in Panel (b) correspond to the sample analogs of equation (25), evaluated at different values of X. We use a Gaussian kernel with a bandwidth of 100,000 NOK, where the evaluation points $\{x\}$ correspond to the cohort-specific deciles of initial income. For each decile, we compute cohort-size weighted averages across cohorts and report equally weighted averages across event times $t = -5$ to -2 and $t = +1$ to $+5$. 95 percent confidence intervals are shown, with standard errors clustered at the winner level. Throughout, we use $g - 1$ as the omitted event time.

income generated by lottery winnings.

Using the same difference-in-differences design as above, we first recover the effect of winning the lottery on unearned income by considering the following estimand,

$$\text{FS}_{g,t}(x) \equiv \underbrace{\mathbb{E}\left[B_{g+t} - B_{g-1} \middle| G = g, Y_{g-1} = x\right]}_{\text{unearned income change for year } g \text{ winners}} - \overbrace{\mathbb{E}\left[B_{g+t} - B_{g-1} \middle| G > g+t, Y_{g-1} = x\right]}^{\text{unearned income change for later-than } g+t\text{-winners}}, \quad (26)$$

for each g, t and x. We estimate and aggregate $\text{FS}_{g,t}(x)$ across cohorts and event times as we did above.

Panel (a) of Figure 6 presents the resulting estimates. It shows no evidence of different pre-trends between winners and later winners in any part of the distribution. Winning the lottery increases unearned income for all income levels. The effect equals around 50,000 for the lowest decile and gradually increases to around 70,000 in the highest decile.

To obtain the earnings response to unearned income, we perform TSLS, taking the ratio of the estimated reduced-form in (25) and first-stage in (26). Panel (b) of Figure 6 reports these IV estimates across the distribution of initial income. It shows that the estimates are remarkably stable across the income distribution. On average, an additional NOK of unearned income reduces earnings by about 0.26, suggesting a marginal propensity to earn out of unearned income of -0.26. For an individual facing a marginal tax rate of 40 percent (corresponding to the sample mean reported in Table A4), this implies a marginal propensity to consume out of unearned income of around 0.84.

5.6 Recovering the intensive-margin income effect

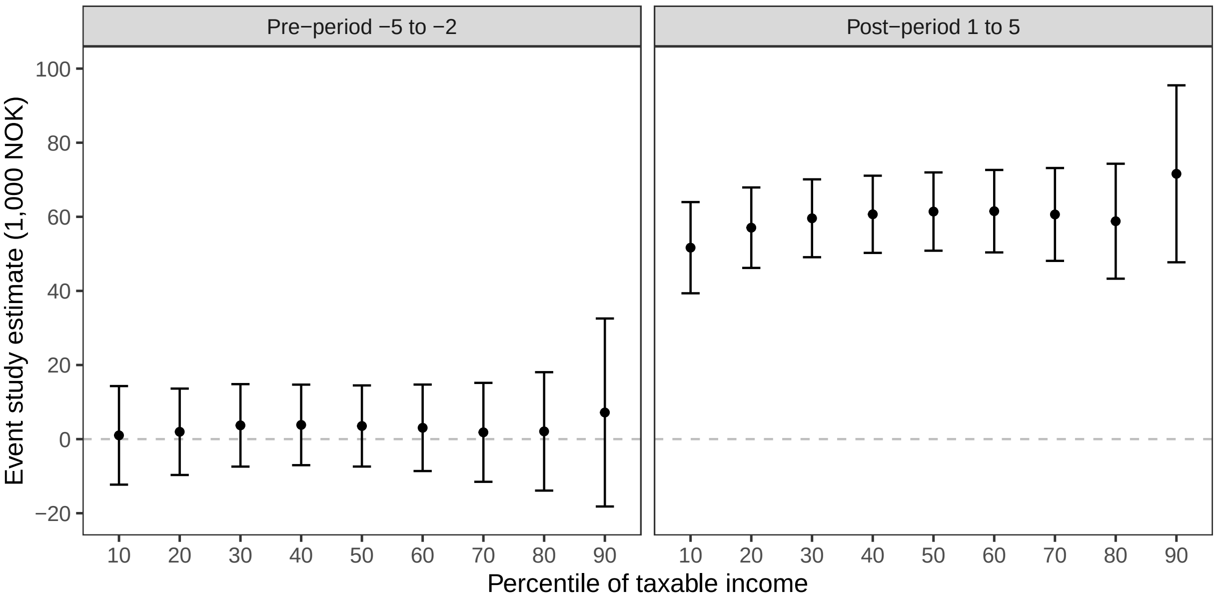
The estimates in Figure 6 reflect a combination of intensive and extensive-margin labor supply responses, while the parameter we need to point identify (un)compensated elasticities ($\eta(x)$) is the intensive-margin labor supply response to changes in unearned income. To recover intensive-margin income effects, we first decompose the total earnings response to unearned income into its intensive- and extensive-margin components. The decomposition infers intensive margin responses for each initial income from the combination of the earnings response and the employment effects by initial income. We refer to Appendix D for a formal derivation of the decomposition.

Panel (a) of Figure 7 plots the results from the decomposition. It shows that the earnings contribution from the intensive margin is by far the most important one, accounting for around 80 percent of the total response. The intensive-margin share increases with income and accounts for nearly 90 percent of the total response for individuals with initial income at the 90th percentile.

To obtain the intensive-margin income response per person, it is useful to note that the intensive-margin contribution equals the share of individuals who respond on the margin multiplied by their response. To obtain these intensive-margin responses, we therefore divide the intensive-margin contribution by the share that responds on the intensive margin, i.e., the share that continues to work after winning.

Panel (b) plots the resulting intensive-margin responses per person. It shows that the intensive-margin earnings response to an additional NOK of unearned income is, on average, -0.22. The estimates vary modestly across the distribution of taxable income, ranging from -0.18 at the 10th

(a) Effect on unearned income



(b) Earnings effect of unearned income

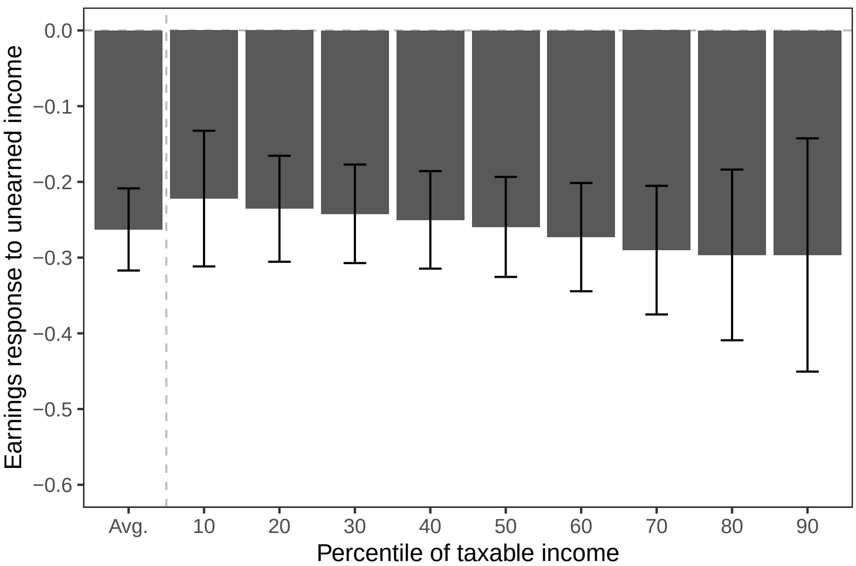


Figure 6: Unearned income effects of winning and IV estimates

Notes: This figure presents estimates of the effect of winning the lottery on winners' unearned income and the earnings response per unearned income. The estimates in Panel (a) correspond to the sample analogs of equation (26), evaluated at different values of X. We use a a Gaussian kernel with a bandwidth of 100,000 NOK, where the evaluation points $\{x\}$ correspond to the cohort-specific deciles of initial income. For each decile, we compute cohort-size weighted averages across cohorts and report equally weighted averages across event times $t = -5$ to -2 and $t = +1$ to $+5$. The estimates in Panel (b) correspond to the ratio between the aggregated values of $\text{RF}_{t,g}(x)$ and $\text{FS}_{t,g}(x)$ evaluated at different values of x. 95 percent confidence intervals are shown, with standard errors clustered at the winner level. Throughout, we use $g - 1$ as the omitted event time.

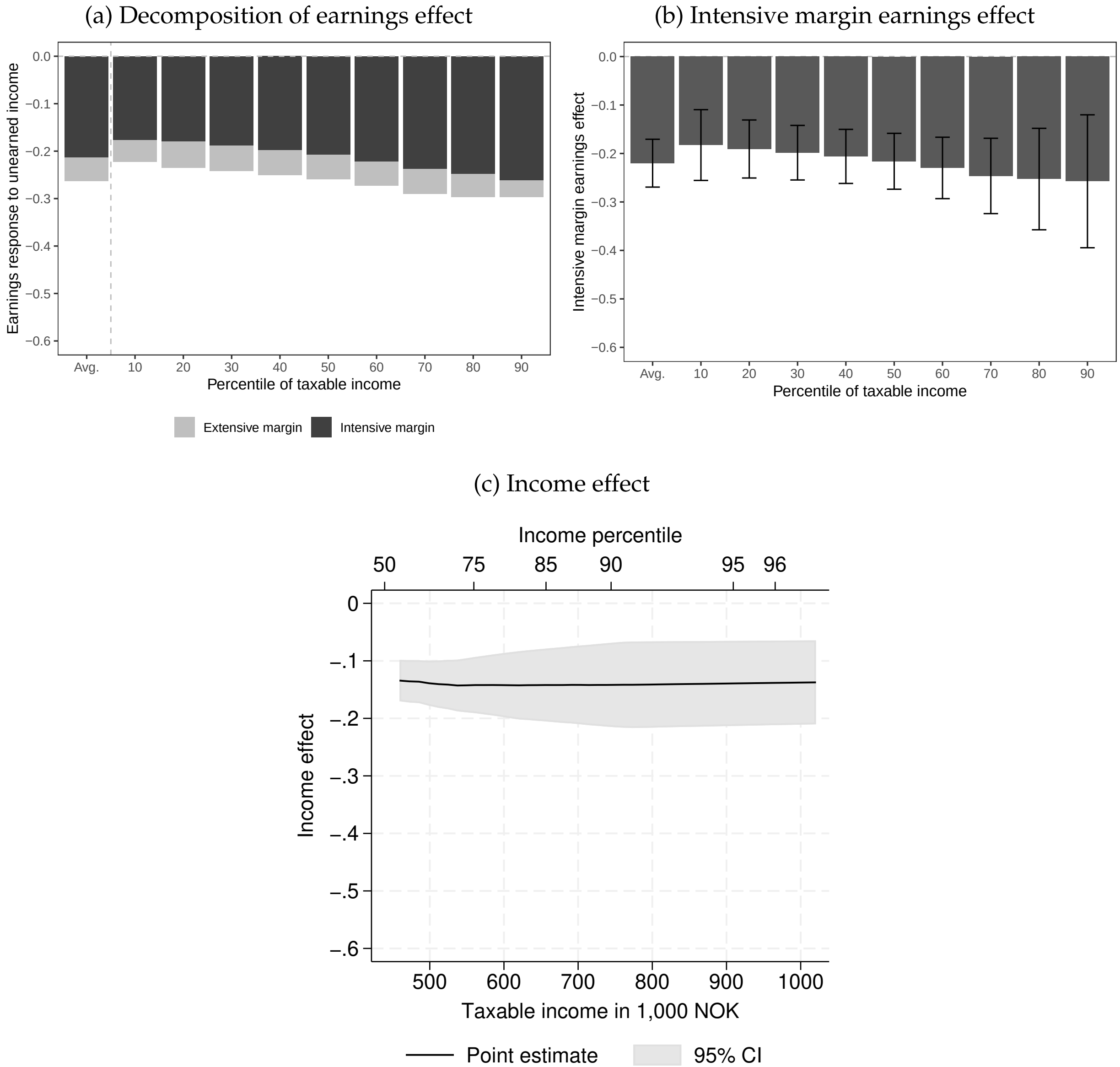


Figure 7: Earnings effect of unearned income and income effects

Notes: Panel (a) decomposes the earnings response per extra NOK of unearned income into its intensive- and extensive-margin components according to Equation (A29). Panel (b) plots the intensive-margin earnings effect conditional on initial income. Panel (c) presents point estimates of income effects $\eta(x)$ that are inferred by multiplying the intensive-margin response reported in Panel (b) by the net-of-tax rate at the corresponding earnings level. 95 percent confidence intervals are shown, with standard errors clustered at the winner level. Throughout, we use $g-1$ as the omitted event time.

percentile to -0.26 at the 90th percentile.

The income effects $\eta(x)$ that we need to point identify the (un)compensated elasticities can now be inferred by multiplying the intensive-margin response reported in Panel (b) by the net-of-tax rate at the corresponding earnings level.[13] These estimates are reported in Panel (c) of Figure 7. The average income effect is -0.14 and does not vary with income.

The estimation sample of the lottery winners is constructed to match the estimation sample for the ETI analysis. In both cases, we include individuals aged 25-61 with wage earnings as their primary labor income source, and exclude students and individuals receiving pensions or unemployment benefits. Nevertheless, one might be concerned that the household and individual characteristics differ between the two samples. We examine this in Table A4 in Appendix B. While the two samples are relatively similar in terms of average initial income, lottery winners tend to be slightly older and less educated than the individuals in the ETI sample.

To address this concern, we construct a new sample consisting of those who won between 1998 and 2003. This ensures that initial income is measured in the same years as in Section 3. We also impose the same restrictions on initial income as we did in Section 3: We exclude winners with initial income below 300,000 or above 1,100,000, and drop those who won before 2001 (corresponding to the placebo cohorts) with initial income between 400,000 and 475,000. Next, we reweight this new lottery sample to match the ETI sample in terms of observable characteristics (including age, gender, marital status, and schooling) and the distribution of initial income.

Table A4 and Figure A3 in Appendix B compare the characteristics and initial income of the original and reweighted lottery sample to the ETI sample. After reweighting, the average individual characteristics and the distribution of initial income of the reweighted lottery sample are virtually indistinguishable from the ETI sample. Figure A4 in Appendix B reports the intensive-margin earnings responses estimated on the reweighted sample. It is reassuring to find that reweighting the sample does not materially change the estimates.

5.7 Point identification with external estimates of income effects

Figure 8 presents point estimates of compensated and uncompensated earnings elasticities obtained using our estimated income effects. It shows that both elasticities are monotonically increasing in income. Panel (a) shows a compensated elasticity of 0.1 for incomes around 450,000 (median) that increases steadily to around 0.45 for incomes close to 750,000 (90th percentile), and equals 0.6 for incomes around 900,000 (95th percentile). On average, the compensated elasticity equals 0.3.

[13] See Appendix F for a derivation.

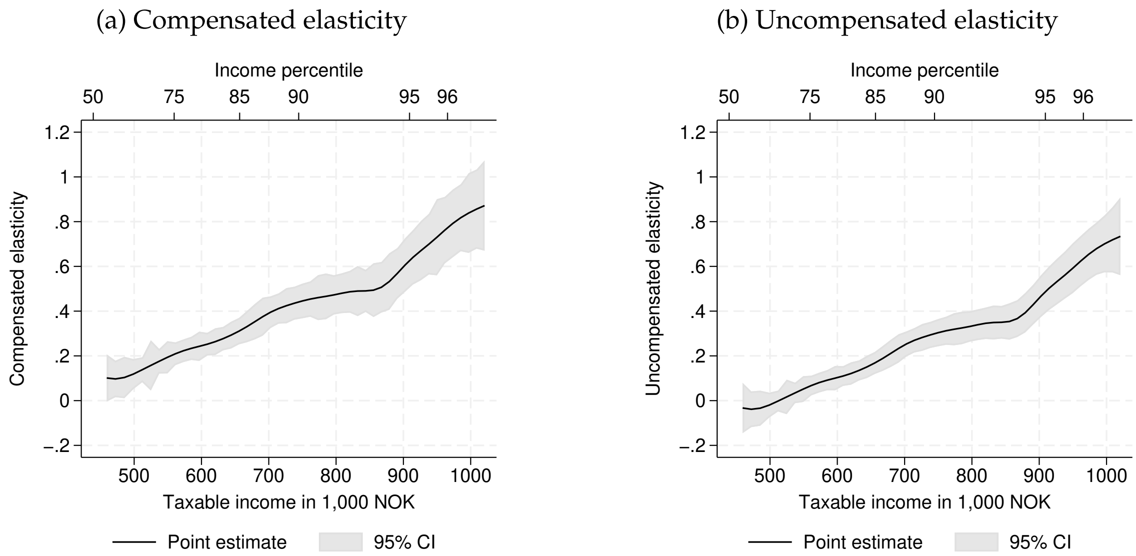


Figure 8: Average (un)compensated earnings elasticities across initial income X.

Notes: The figure plots point estimates of the compensated (Panel (a)) and uncompensated (Panel (b)) labor supply elasticities over the income distribution. 95 percent confidence intervals are shown, with standard errors obtained by assuming that the lottery-based income effects $\eta(x)$ and local ETIs $\beta(x)$ are uncorrelated.

Panel (b) shows that the uncompensated elasticity is close to zero for individuals with income between 450,000 and 500,000. For incomes above the 75th percentile, the uncompensated elasticity is positive and significantly different from zero. It equals 0.3 for incomes around 750,000 (90th percentile) and increases steadily to around 0.5 for incomes around 900,000 NOK (95th percentile). On average, the uncompensated elasticity equals 0.16. This suggests substitution effects are larger than income effects, especially for individuals with relatively high incomes.

6 Revenue-maximizing tax rates and excess burden

In this section, we use our estimates of income and substitution effects to quantify the efficiency cost and revenue effect of increasing the marginal tax rates on middle- and high-income individuals. We also calculate the implied revenue-maximizing top-income tax rate. To do so, we begin by tailoring the standard expressions for marginal deadweight loss (Auerbach and Hines, 2002 and Harberger, 1964), the revenue function (Bierbrauer et al., 2023), and the revenue-maximizing top income tax rate (Diamond, 1998 and Saez, 2001) to the Norwegian institutional context by accounting for payroll and value-added taxes. Next, we evaluate these expressions using our estimates of income and substitution effects from Section 5.

6.1 Measuring the effective tax rate

To understand how changes in marginal tax rates affect economic efficiency and tax revenue, it is necessary to account for the value-added and payroll taxes that are part of the Norwegian tax system. To this end, we suppose the tax revenue collected from an individual with a consumption bundle $(C_1, ..., C_K)$ and earnings Y equals,

$$TR(C_1, ..., C_K, Y) \equiv \underbrace{\sum_{k=1}^{K} \tau_k P_k C_k}_{\text{consumption tax}} + \overbrace{T(Y)}^{\text{income tax}} + \underbrace{\tau_w Y}_{\text{payroll tax}},$$

where τ_w is the proportional payroll tax rate and τ_k is the tax rate on consumption good k.

The tax rates τ_k vary across consumption goods. This means that to calculate the revenue effect of changes in the income tax, one would need to know how it affects the demand for each of the K consumption goods. To reduce the dimensionality of the problem, we specialize our model from Section 4 by assuming that individual preferences are separable between the consumption and labor supply components:

$$U(C_1, ..., C_K, Y_1, ..., Y_L) = \underbrace{U_c(C_1, ..., C_K)}_{\text{utility from consumption}} + \overbrace{U_y(Y_1, ..., Y_L)}^{\text{disutility from labor}}.$$

By standard two-stage budgeting arguments, this means that the uncompensated demand for the k-th consumption good depends only on the income tax through its effect on disposable income:

$$C_k = C_k^u\Big(\tau_1, ..., \tau_K, \underbrace{I^u(\tau_1, ..., \tau_K, \tau, R + B)}_{\equiv(1-\tau)Y^u(\tau_1,...,\tau_K,\tau,R+B)+R+B}\Big). \tag{27}$$

Using these demand functions, we can express the effective consumption tax rate as

$$\tilde{\tau} \equiv \frac{\sum_{k=1}^{K} \tau_k P_k \frac{\partial C_k^u}{\partial I}}{\sum_{k=1}^{K} P_k \frac{\partial C_k^u}{\partial I}}. \tag{28}$$

As we show below, this tax rate is a sufficient statistic for calculating how changes in disposable income affect consumption tax revenue.

Assuming that each consumption good is normal ensures that $\tilde{\tau}$ is bounded between the lowest and largest consumption tax rate. In our setting, this means that the effective consumption tax rate is bounded between 0 and 0.25. We allow for any effective consumption tax rate in this range but impose that it does not vary across individuals.[14] Using average expenditure shares to proxy for the marginal expenditure shares implies that $\tilde{\tau} = 0.21$.

6.2 Excess burden of taxation

We now quantify the efficiency cost, measured by the excess burden, of increasing the marginal tax rate for middle- and high-income individuals. It is useful to start by considering the individual's

[14] This is true if, for example, the consumption component of the utility function U_c is homothetic and homogeneous across individuals.

expenditure minimization problem:

$$E(\tau_1, ..., \tau_K, \tau, \bar{V}) \equiv \min_{C_1,...,C_K,Y_1,...,Y_L} \sum_{k=1} (1+\tau_k)P_k C_k - (1-\tau)F(Y_1, ..., Y_L) \quad \text{subject to } U(C_1, ..., C_K, Y_1, ..., Y_L) \geq \bar{V}, \tag{29}$$

where $\bar{V}$ is individual utility after the tax reform. The resulting compensated demand functions can be written as $C_k^c(\tau_1, ..., \tau_K, \tau, \bar{V})$ and the earnings function as $Y^c(\tau_1, ..., \tau_K, \tau, \bar{V}) \equiv F(Y_1^c(\tau_1, ..., \tau_K, \tau, \bar{V}), ..., Y$

These functions allow us to define the deadweight loss of taxation,

$$DWL(\tau_1, ..., \tau_K, \tau, \bar{V}) \equiv E(\tau_1, ..., \tau_K, \tau, \bar{V}) - TR^c(\tau_1, ..., \tau_K, \tau, \bar{V}),$$

where the compensated tax revenue TR^c is given by,

$$TR^c(\tau_1, ..., \tau_K, \tau, \bar{V}) \equiv \sum_{k=1}^{K} \tau_k P_k C_k^c(\tau_1, ..., \tau_K, \tau, \bar{V}) + (\tau + \tau_w)Y^c(\tau_1, ..., \tau_K, \tau, \bar{V}).$$

In words, the deadweight loss is the additional expenditure, beyond the revenue raised, that is required to keep an individual at utility level $\bar{V}$ under the tax system $(\tau_1, ..., \tau_K, \tau)$.

The marginal deadweight loss is obtained by taking the derivative of DWL with respect to τ. The excess burden EB normalizes the marginal deadweight loss by the marginal revenue raised, i.e., the derivative of TR^c with respect to τ:

$$EB \equiv \frac{\mathbb{E}\left[\frac{\partial DWL}{\partial \tau} \mid Y \geq \bar{Y}\right]}{\mathbb{E}\left[\frac{\partial TR^c}{\partial \tau} \mid Y \geq \bar{Y}\right]}. \tag{30}$$

In words, the excess burden EB measures the economic cost of a marginal increase in the marginal tax rate per additional NOK of revenue raised.

We show in Appendix E.1 that the excess burden can be expressed in terms of compensated elasticities, tax rates, and earnings as:

$$EB = \frac{\mathbb{E}\left[\left(\frac{\tilde{\tau}}{1+\tilde{\tau}} + \frac{\tau+\tau_w}{1-\tau}\right)\varepsilon^c Y \mid Y \geq \bar{Y}\right]}{\mathbb{E}\left[\left(1 - \left(\frac{\tilde{\tau}}{1+\tilde{\tau}} + \frac{\tau+\tau_w}{1-\tau}\right)\varepsilon^c\right) Y \mid Y \geq \bar{Y}\right]}. \tag{31}$$

We now calculate the excess burden in equation (31) under two different sets of elasticity estimates. The first set is obtained from our data and estimates of ε^c from Section 5, reported in Panel (a) of Figure 8. We refer to these estimates as our preferred specification. The second is the compensated elasticities one would obtain from using the common ETI estimand, ignoring income effects, and assuming constant labor supply elasticities. Under this conventional specification, $\varepsilon^c = 0.16$ and $\eta = 0$ for all individuals.[15]

Panel (a) of Figure 9 plots the excess burden implied by the two sets of estimates as functions of the effective consumption tax rate $\tilde{\tau}$. The solid black line shows the excess burden using our preferred specification. It shows that increasing the marginal tax rate on middle- and high-income individuals results in an economic loss of at least 0.85 NOK for every additional NOK

[15] To see why, note that under homogeneous ETIs and no income effects, Corollary 1 and Proposition 3 together imply that the ETI estimand with $h(Z) = Z$ estimates the compensated elasticity. Hence $\beta = 0.16 = \zeta = \varepsilon^c$.

of tax revenue raised. For $\tilde{\tau} = 0.21$, based on observed expenditure shares, the excess burden is approximately 1.1.

This contrasts sharply with the results obtained using the conventional specification, as depicted by the dashed grey line. The conventional specification implies a more modest excess burden of 0.23-0.28, depending on the exact value of the effective tax rate $\tilde{\tau}$. We conclude that ignoring income effects and heterogeneity in elasticities can lead to severely downward-biased estimates of excess burden. In our setting, the actual efficiency costs of taxation are four to five times larger than those implied by the conventional specification.

6.3 Pareto-improving local tax reforms

Bierbrauer et al. (2023) shows that the *revenue function*, defined as the effect of raising marginal tax rates in a small bracket at a certain income level on tax revenue, can be used to test if there exists local Pareto-improving tax reforms. In our model, the revenue function is given by

$$\mathcal{R}(y) = \frac{1 - F(y) - f(y)\tau(y)\frac{\varepsilon^c(y)y}{1-T'(y)} - \int_y^{+\infty} f(z)\tau(z)\frac{\eta(z)}{1-T'(z)}dz}{1 - \int_0^{+\infty} f(z)\tau(z)\frac{\eta(z)}{1-T'(z)}dz}, \tag{32}$$

with

$$\tau(y) \equiv T'(y) + \tau_w + \frac{\tilde{\tau}}{1+\tilde{\tau}}(1 - T'(y)).$$

$\tau(y)$ measures the marginal effect of changes in earnings on tax revenue, T is the current tax system, F is the cumulative distribution function of earnings Y, and f is the corresponding probability density function. Proposition 2 of Bierbrauer et al. (2023) states that the tax system is locally Pareto-efficient if and only if $\mathcal{R}$ is non-increasing and bounded between 0 and 1.

Panel (b) of Figure 9 plots the revenue function $\mathcal{R}$ with $\tilde{\tau} = 0.21$ over the relevant range of incomes. The solid line is calculated using our data and estimates of $\varepsilon^c(x)$ and $\eta(x)$ from Section 5. We refer to these estimates as our preferred specification. It shows that the revenue function is decreasing and bounded between 0 and 1 for incomes below 650,000, implying that no Pareto-improving tax reform is available over this interval. For incomes above 650,000, the revenue function is negative, indicating that reducing marginal tax rates for these incomes would increase tax revenue. Thus, a tax reform that reduces the marginal tax rate on incomes above 650,000 and rebates the additional revenue collected lump sum would increase welfare for all individuals.

The conventional specification with $\varepsilon^c = 0.16$ and $\eta = 0$ yields a revenue function that underestimates (overestimates) the revenue effects of raising marginal tax rates for incomes below (above) 550,000. This means that it underestimates the cost of making the marginal tax rate schedule steeper. More importantly, because the function is decreasing and bounded between 0 and 1 over the entire income range, it would fail to identify any Pareto-improving tax reform.

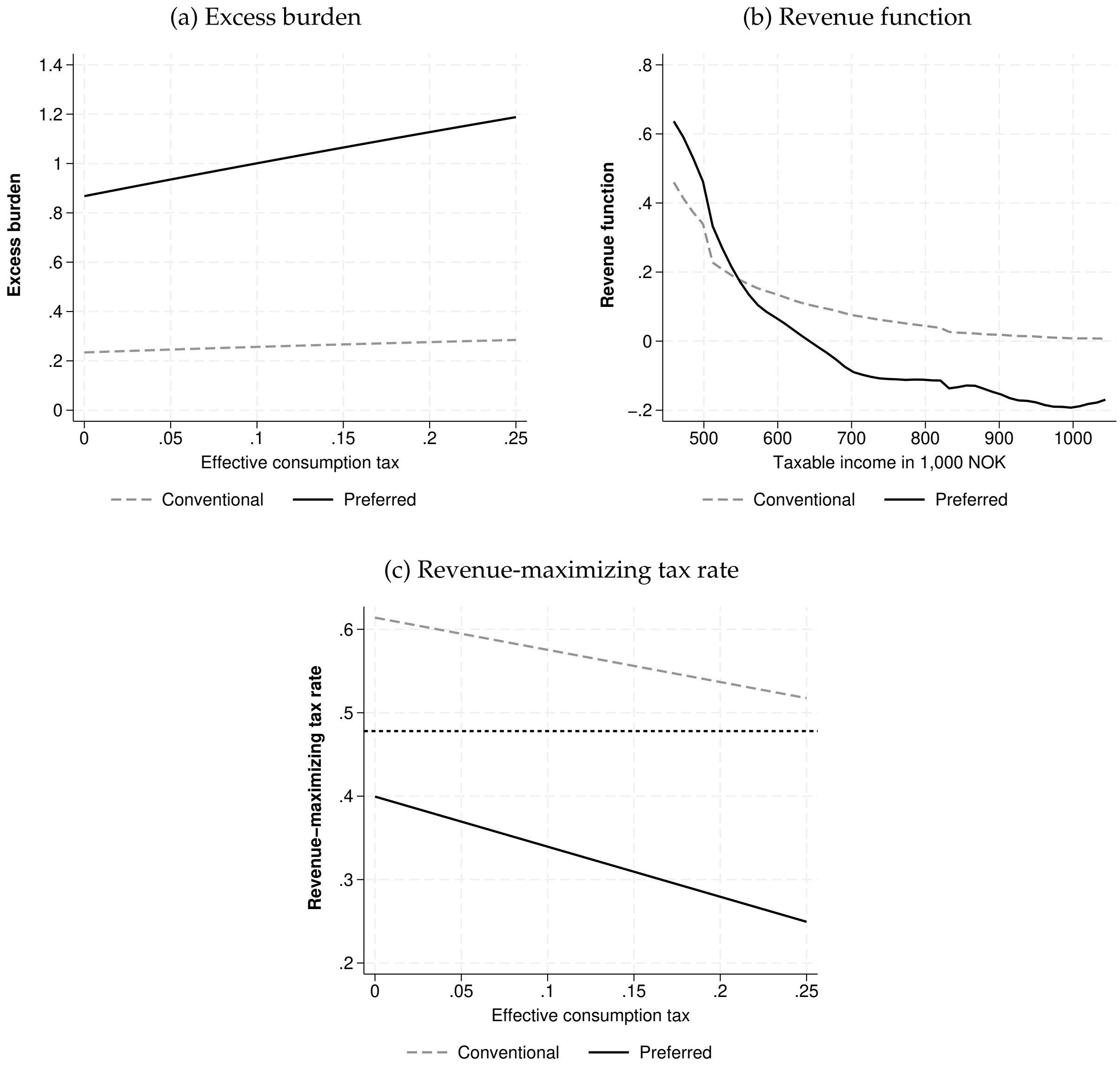


Figure 9: Excess burden, revenue functions and revenue-maximizing top-income tax rates.

Notes: Panel (a) plots the excess burden of taxation as a function of $\tilde{\tau}$ using our preferred estimates (the solid black line) and using the estimates one would obtain from using the common ETI estimand, ignoring income effects, and assuming constant labor supply elasticities (dashed grey line). Panel (b) plots the revenue function in equation (32) with $\tilde{\tau} = 0.21$ using the same two sets of estimates. Panel (c) plots the revenue-maximizing top-income tax rates as a function of $\tilde{\tau}$ using the same two sets of estimates. The dotted line shows the actual top-income tax rate in Norway after the 2006 reform.

6.4 Revenue-maximizing tax rate

We now turn to deriving and computing the revenue-maximizing top-income tax rate. It is useful to start by considering the tax revenue collected from individuals with earnings above $\bar{Y}$,

$$\mathbb{E}[TR \mid Y \geq \bar{Y}] = \overbrace{\sum_{k=1}^{K} \tau_k p_k \, \mathbb{E}[C_k \mid Y \geq \bar{Y}]}^{\text{consumption tax}} + \overbrace{\tau \, \mathbb{E}[Y - \bar{Y} \mid Y \geq \bar{Y}] + T(\bar{Y})}^{\text{income tax}} + \underbrace{\tau_w \, \mathbb{E}[Y \mid Y \geq \bar{Y}]}_{\text{payroll tax}}, \tag{33}$$

where the expectations are taken across individuals in the top-income tax bracket.

We characterize the revenue-maximizing top-income tax rate by taking the derivative of equation (33) with respect to the top-income tax rate τ and setting it equal to zero. As we show in Appendix E.2, this first-order condition implies that the following equation is satisfied,

$$\tau_{\text{TOP}} = \frac{\alpha - 1 - (\tilde{\tau} + \tau_w + \tau_w \tilde{\tau}) \, (\alpha \bar{\varepsilon}^u - \bar{\eta})}{\alpha - 1 + (\alpha \bar{\varepsilon}^u - \bar{\eta})}, \tag{34}$$

with

$$\bar{\varepsilon}^u \equiv \frac{\mathbb{E}[Y^u \varepsilon^u \mid Y^u \geq \bar{Y}]}{\mathbb{E}[Y^u \mid Y^u \geq \bar{Y}],} \qquad \bar{\eta} \equiv \mathbb{E}[\eta \mid Y^u \geq \bar{Y}], \qquad \alpha \equiv \frac{\mathbb{E}[Y^u \mid Y^u \geq \bar{Y}]}{\bar{Y}}, \tag{35}$$

where $Y^u, \varepsilon^u, \eta, \alpha$, and $\tilde{\tau}$ are evaluated at the revenue-maximizing tax system.

We consider the tax rate that maximizes tax revenue on incomes above $\bar{Y} = 450{,}000$. Following Saez (2001) and Saez and Stantcheva (2016), we assume that the weighted elasticities $\bar{\varepsilon}^u$, $\bar{\eta}$, and the Pareto parameter α are unaffected by the top tax rate. Using the estimates and data in Section 5, we obtain $\alpha = 1.46$, $\bar{\varepsilon}^u = 0.25$, and $\bar{\eta} = -0.14$. As above, we refer to this set of elasticities as our preferred specification. For comparison, we also calculate the revenue-maximizing top-income tax rate under the conventional specification, obtained by using the common ETI estimand, ignoring income effects, and assuming constant labor supply elasticities. Under this specification, $\bar{\varepsilon}^u = 0.16$ and $\bar{\eta} = 0$.

Panel (c) of Figure 9 plots the revenue-maximizing top-income tax rates implied by these two specifications as functions of the effective consumption tax rate $\tilde{\tau}$. The solid black line shows the results using our preferred specification. The revenue-maximizing top rate is at most 0.4. It declines with the effective consumption tax rate $\tilde{\tau}$, and reaches 0.25 when $\tilde{\tau}$ is at its upper bound. For $\tilde{\tau} = 0.21$, obtained using observed expenditure shares, the revenue-maximizing tax rate is around 0.27. The gray dashed line uses the elasticities implied by the conventional specification. The elasticities implied by the conventional specification increase the revenue-maximizing tax rate by more than 20 percentage points.

Interestingly, our estimates imply that the actual top-income tax rate after the 2006 reform, illustrated by the dotted black line, exceeds the revenue-maximizing level for any effective consumption tax rate $\tilde{\tau}$. This means that the government could increase revenue by *reducing* top-income tax rates.

7 Implications for Frisch elasticities and EIS

We now consider the implications of our findings for the Frisch elasticity and the elasticity of intertemporal substitution (EIS) by deriving the relationship between these quantities and the compensated elasticity and income effect estimates in Section 5.

We consider the textbook intertemporal problem where individuals choose consumption and earnings to maximize the discounted sum of per-period utility subject to the intertemporal budget constraint in equation (24). Following MaCurdy (1981), the per-period utility function is given by:

$$U(C, Y) = \frac{C^{1-1/\rho}}{1 - 1/\rho} - \frac{\Psi}{1 + 1/\varepsilon^f} \left(\frac{Y}{\Psi}\right)^{1+1/\varepsilon^f}, \tag{36}$$

which is a special case of the utility function we considered in Section 4. The EIS is then given by ρ, and the Frisch elasticity is equal to ε^f.[16]

It follows from Keane (2011) that ε^f and ρ are linked to the compensated elasticity and income effect through the following relationships:

$$\varepsilon^f = \frac{\varepsilon^c}{1 + \eta}, \qquad \rho = -(1 - \tau)\frac{Y_t}{C_t}\frac{\varepsilon^c}{\eta}, \tag{37}$$

where τ is the marginal tax rate and ε^c and η are defined as in equation (16) in Section 4. These expressions show how our estimates of compensated elasticities and income effects can be used to recover the EIS and the Frisch elasticity. At each level of initial income X, the average EIS and Frisch elasticity are given by:

$$\begin{aligned} \mathbb{E}[\varepsilon^f \mid X = x] &= \frac{\varepsilon^c(x)}{1 + \eta(x)}, \\ \mathbb{E}[\rho \mid X = x] &= -\,\mathbb{E}\left[(1 - T'(Y_t))\frac{Y_t}{C_t} \mid X = x\right] \times \frac{\varepsilon^c(x)}{\eta(x)}, \end{aligned}$$

where each component of the RHS can be estimated in the data.

Using our estimated compensated elasticities and income effects from Section 5, we find an average Frisch elasticity of 0.35. This finding is slightly lower than the average point estimate of 0.5 reported in the meta-analysis of Elminejad et al. (2023). However, as shown in Panel (a) of Figure 10, this average masks considerable heterogeneity. The Frisch elasticity equals 0.1 for incomes around 450,000 (median). It increases steadily to around 0.5 for incomes close to 750,000 (90th percentile), and equals 0.7 for incomes around 900,000 (95th percentile).

In Panel (b) of Figure 10, we report the EIS estimates across the income distribution. On average, the EIS is about 1.58. This finding aligns closely with Holm et al.'s (2024) constant effect estimates of the EIS in Norway. However, there is systematic heterogeneity in the EIS across the income distribution. The EIS equals 0.5 for individuals with an income of around 450,000 (median), increases steadily to around 2.5 for incomes close to 750,000 (90th percentile), and is about 4 for incomes around 900,000 (95th percentile).

[16] Under the utility function in (36), the EIS is inversely related to the coefficient of relative risk aversion. See Chetty (2006) for how labor supply elasticities can be used to learn about risk preferences.

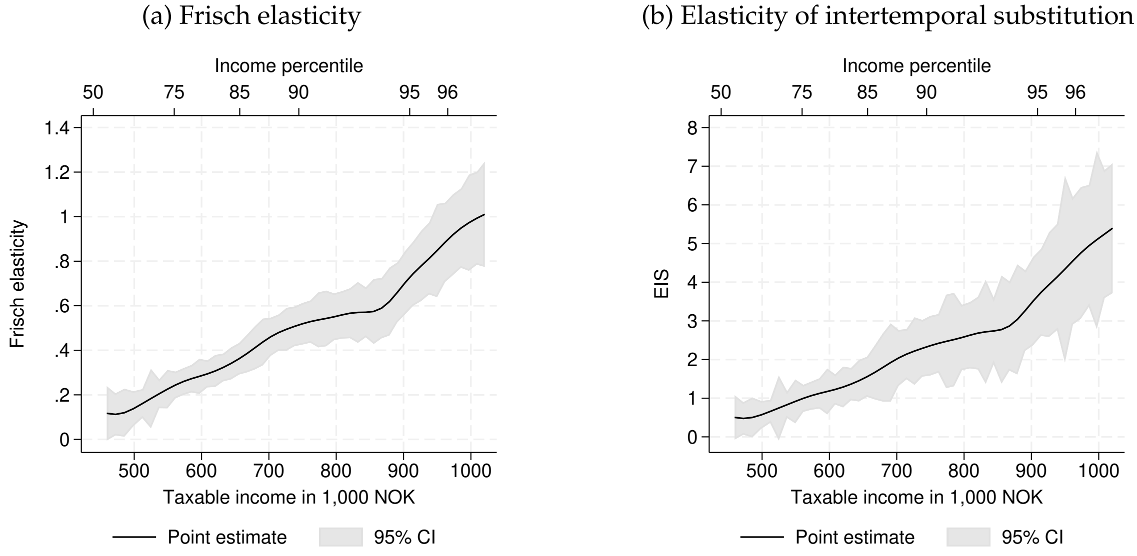


Figure 10: Frisch elasticities and EIS.

Notes: The figure plots point estimates of the Frisch elasticities (Panel (a)) and consumption-weighted EIS (Panel (b)) over the income distribution. 95 percent confidence intervals are shown, with standard errors obtained using the delta method, assuming that the lottery-based income effects $\eta(x)$ and local ETIs $\beta(x)$ are uncorrelated.

8 Conclusions

This paper examined when a commonly applied class of ETI estimands can be used to learn about individuals' ETI parameters and their (un)compensated elasticities of labor supply. We provided necessary and sufficient conditions for these estimands to be given a causal interpretation as a positively weighted average of heterogeneous ETI parameters. We then applied these results to empirically analyze a reform of the Norwegian tax system that reduced the marginal tax rates on middle and high incomes. The estimated ETI parameters increase steadily with income, meaning high-income individuals are more responsive to tax changes than middle-income individuals.

Next, we showed how (un)compensated elasticities of labor supply can be bounded directly from the ETI estimands, or point identified by combining these estimands with estimates of earnings responses to lottery winnings. The results suggest an (un)compensated elasticity of 0.1 (0.0) for middle-income individuals. The (un)compensated elasticity estimates increase steadily with income to around 0.45 (0.3) for high-income individuals. These findings imply a substantial excess burden of taxation, and that reducing the top-income tax rate would increase tax revenue. Our findings are also informative about how the intertemporal elasticity of substitution and the Frisch elasticity vary across the income distribution.

References

Aaberge, R., & Colombino, U. (2018). Structural labour supply models and microsimulation. *International Journal of Microsimulation*, *11*(1), 162–197.

Altonji, J. G. (1986). Intertemporal substitution in labor supply: Evidence from micro data. *Journal of Political Economy*, *94*(3, Part 2), S176–S215.

Angrist, J. D. (1998). Estimating the labor market impact of voluntary military service using social security data on military applicants. *Econometrica*, *66*(2), 249–288.

Angrist, J. D., & Imbens, G. W. (1995). Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American statistical Association*, *90*(430), 431–442.

Arellano, M., & Meghir, C. (1992). Female labour supply and on-the-job search: An empirical model estimated using complementary data sets. *The Review of Economic Studies*, *59*(3), 537–559.

Attanasio, O., Levell, P., Low, H., & Sánchez-Marcos, V. (2018). Aggregating elasticities: Intensive and extensive margins of women's labor supply. *Econometrica*, *86*(6), 2049–2082.

Attanasio, O. P., & Weber, G. (1995). Is consumption growth consistent with intertemporal optimization? evidence from the consumer expenditure survey. *Journal of Political Economy*, *103*(6), 1121–1157.

Auerbach, A. J., & Hines, J. R. J. (2002). Taxation and economic efficiency. In A. J. Auerbach & M. Feldstein (Eds.), *Handbook of public economics* (pp. 1347–1421, Vol. 3). Elsevier.

Auten, G., & Carroll, R. (1999). The effect of income taxes on household income. *The Review of Economics and Statistics*, *81*(4), 681–693.

Berg, K., & Thoresen, T. O. (2020). Problematic response margins in the estimation of the elasticity of taxable income. *International Tax and Public Finance*, *27*(3), 721–752.

Bierbrauer, F. J., Boyer, P. C., & Hansen, E. (2023). Pareto-improving tax reforms and the earned income tax credit. *Econometrica*, *91*(3), 1077–1103.

Blandhol, C., Bonney, J., Mogstad, M., & Torgovitsky, A. (2022, January). *When is tsls actually late?* (Working Paper No. 29709). National Bureau of Economic Research.

Blomquist, N. S. (1985). Labour supply in a two-period model: The effect of a nonlinear progressive income tax. *The Review of Economic Studies*, *52*(3), 515–524.

Blundell, R., & Macurdy, T. (1999). Chapter 27 - labor supply: A review of alternative approaches. In O. C. Ashenfelter & D. Card (Eds.). Elsevier.

Blundell, R., & Walker, I. (1986). A life-cycle consistent empirical model of family labour supply using cross-section data. *The Review of Economic Studies*, *53*(4), 539–558.

Browning, M., & Lusardi, A. (1996). Household saving: Micro theories and micro facts. *Journal of Economic Literature*, *34*(4), 1797–1855.

Bulman, G., Cesarini, D., Lindqvist, E., Notowidigdo, M. J., & Östling, R. (2021). The impact of lottery prizes on winners and their neighbors: Evidence from the dutch postcode lottery. *American Economic Review: Insights*, *3*(3), 333–350.

Burns, S. K., & Ziliak, J. P. (2017). Identifying the elasticity of taxable income. *Economic Journal*, *127*(600), 297–329.

Callaway, B., Goodman-Bacon, A., & Sant'Anna, P. H. C. (2025). Difference-in-differences with a continuous treatment.

Cesarini, D., Lindqvist, E., Notowidigdo, M. J., & Östling, R. (2017). The effect of wealth on individual and household labor supply: Evidence from swedish lotteries. *American Economic Review*, *107*(12), 3917–3946.

Chetty, R. (2006). A new method of estimating risk aversion. *American Economic Review*, *96*(5), 1821–1834.

Chetty, R., Guren, A., Manoli, D., & Weber, A. (2011). Are micro and macro labor supply elasticities consistent? A review of evidence on the intensive and extensive margins. *American Economic Review: Papers & Proceedings*, *101*(3), 471–475.

Conesa, J. C., Kitao, S., & Krueger, D. (2009). Taxing capital? not a bad idea after all! *American Economic Review*, *99*(1), 25–48.

Diamond, P. A. (1998). Optimal income taxation: An example with a u-shaped pattern of optimal marginal tax rates. *The American Economic Review*, *88*(1), 83–95.

Eika, L., Mogstad, M., & Vestad, O. L. (2020). What can we learn about household consumption expenditure from data on income and assets? *Journal of Public Economics*, *189*, 104163.

Elminejad, A., Havránek, T., & Irsová, Z. (2023). Intertemporal substitution in labor supply: A meta-analysis. *Labour Economics*, *83*, 102404.

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.

Frisch, R., & Waugh, F. V. (1933). Partial time regressions as compared with individual trends. *Econometrica*, *1*(4), 387–401.

Golosov, M., Graber, M., Mogstad, M., & Novgorodsky, D. (2024). How americans respond to idiosyncratic and exogenous changes in household wealth and unearned income. *The Quarterly Journal of Economics*, *139*(2), 1321–1395.

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

Hall, R. E. (1973). Wages, income and hours of work in the us labor force. *Income Maintenance and Labor Supply: Econometric Studies (Rand McNally, Chicago)*.

Hall, R. E. (1988). Intertemporal substitution in consumption. *Journal of Political Economy*, *96*(2), 339–357.

Hansen, L. P., & Singleton, K. J. (1983). Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *Journal of Political Economy*, *91*(2), 249–265.

Harberger, A. (1964). Taxation, resource allocation, and welfare. In *The role of direct and indirect taxes in the federal reserve system* (pp. 25–80). National Bureau of Economic Research, Inc.

Havránek, T. (2015). Measuring intertemporal substitution: The effect of different data. *Journal of the European Economic Association*, *13*(6), 1180–1204.

Holm, M. B., Jamilov, R., Jasinski, M., & Nenov, P. T. (2024). *Estimating the elasticity of intertemporal substitution using dividend tax news shocks* (tech. rep. No. 17/2024). Norges Bank Working Paper.

Horowitz, J. L., & Manski, C. F. (1995). Identification and robustness with contaminated and corrupted data. *Econometrica: Journal of the Econometric Society*, 281–302.

Imbens, G. W., & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, *62*(2), 467–475.

Imbens, G. W., Rubin, D. B., & Sacerdote, B. I. (2001). Estimating the effect of unearned income on labor earnings, savings, and consumption: Evidence from a survey of lottery players. *American Economic Review*, *91*(4), 778–794.

Keane, M., & Neal, T. (2025). Robust inference for the frisch labor supply elasticity. *Journal of Labor Economics*, *43*(S1), S179–S219.

Keane, M. P. (2011). Labor supply and taxes: A survey. *Journal of Economic Literature*, *49*(4), 961–1075.

Killingsworth, M. R., & Heckman, J. J. (1986). Female labor supply: A survey. In O. Ashenfelter & R. Layard (Eds.), *Handbook of labor economics* (pp. 103–204, Vol. 1). Elsevier.

Kleven, H. J., & Schultz, E. A. (2014). Estimating taxable income responses using danish tax reforms. *American Economic Journal: Economic Policy*, *6*(4), 271–301.

Kopczuk, W. (2005). Tax bases, tax rates and the elasticity of reported income. *Journal of Public Economics*, *89*(11), 2093–2119.

Kumar, A., & Liang, C.-Y. (2020). Estimating taxable income responses with elasticity heterogeneity. *Journal of Public Economics*, *188*, 104209.

Lee, D. S. (2005). Training, wages, and sample selection: Estimating sharp bounds on treatment effects.

Lovell, M. C. (1963). Seasonal adjustment of economic time series and multiple regression analysis. *Journal of the American Statistical Association*, *58*(304), 993–1010.

MaCurdy, T. E. (1981). An empirical model of labor supply in a life-cycle setting. *Journal of Political Economy*, *89*(6), 1059–1085.

MaCurdy, T. E. (1983). A simple scheme for estimating an intertemporal model of labor supply and consumption in the presence of taxes and uncertainty. *International Economic Review*, *24*(2), 265–289.

Mogstad, M., Salvanes, K. G., & Torsvik, G. (2025). Income equality in the nordic countries: Myths, facts, and lessons. *Journal of Economic Literature*, *63*(3), 791–839.

Mogstad, M., & Torgovitsky, A. (2024). Chapter 1 - instrumental variables with unobserved heterogeneity in treatment effects. In C. Dustmann & T. Lemieux (Eds.). Elsevier.

Neisser, C. (2021). The elasticity of taxable income: A meta-regression analysis. *The Economic Journal*, *131*(640), 3365–3391.

Pencavel, J. (1986). Labor supply of men: A survey (O. Ashenfelter & R. Layard, Eds.). *Handbook of Labor Economics*, *1*, 3–102.

Picchio, M., Suetens, S., & van Ours, J. C. (2018). Labor supply effects of winning a lottery. *Economic Journal*, *128*(610), 1700–1729.

Prescott, E. C. (2006). Nobel lecture: The transformation of macroeconomic policy and research. *Journal of Political Economy*, *114*(2), 203–235.

Saez, E. (2001). Using Elasticities to Derive Optimal Income Tax Rates. *The Review of Economic Studies*, *68*(1), 205–229.

Saez, E., Slemrod, J., & Giertz, S. H. (2012). The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of economic literature*, *50*(1), 3–50.

Saez, E., & Stantcheva, S. (2016). Generalized social marginal welfare weights for optimal tax theory. *American Economic Review*, *106*(1), 24–45.

Straub, L., & Werning, I. (2020). Positive long-run capital taxation: Chamley-judd revisited. *American Economic Review*, *110*(1), 86–119.

Thoresen, T. O., & Vattø, T. E. (2015). Validation of the discrete choice labor supply model by methods of the new tax responsiveness literature. *Labour Economics*, *37*, 38–53.

Vattø, T. E. (2020). Estimating the elasticity of taxable income when earnings responses are sluggish. *FinanzArchiv / Public Finance Analysis*, *76*(4), 329–369.
Weber, C. E. (2014). Toward obtaining a consistent estimate of the elasticity of taxable income using difference-in-differences. *Journal of Public Economics*, *117*, 90–103.

A Proofs

A.1 Proofs of results from Section 2

The following two lemmas will be useful for many of our subsequent results. The first expresses the ETI estimand under Assumptions 1 and 2 for arbitrary h and f functions. The second demonstrates that the condition on f in equation (10) is necessary and sufficient to ensure that $\mathbb{E}[\bar{f}(X)d(G,X)] = 0$ for any arbitrary function $\bar{f}$.

Lemma 1. *Suppose Assumptions 1 and 2 hold. Then, the ETI estimand β equals*

$$\beta = \frac{\mathbb{E}[f^{y}(X)d(G,X)] + \mathbb{P}(G=1)\,\mathbb{E}[\mathbb{1}[X \geq \bar{Y}]d(G,X)\Phi\zeta \mid G=1]}{\mathbb{E}[f^{\text{NTR}}(X)d(G,X)] + \mathbb{P}(G=1)\,\mathbb{E}[\mathbb{1}[X \geq \bar{Y}]d(G,X)\Phi \mid G=1]}.$$

where $d(G,X) \equiv G\bar{h}(X) - \mathbb{L}[Gh(Z) \mid G,X]$ and $\bar{h}$ is defined as

$$\bar{h}(X) \equiv h\left(\log\left(\frac{1-T_1'(X)}{1-T_0'(X)}\right)\right).$$

Proof of Lemma 1. By definition, $h(Z) = \bar{h}(X)$ with probability one. Thus, the Frisch-Waugh-Lovell theorem (Frisch and Waugh, 1933; Lovell, 1963) implies that the ETI estimand can be written as,

$$\beta \equiv \frac{\mathbb{E}\left[\Delta y \times d(G,X)\right]}{\mathbb{E}\left[\Delta\,\text{NTR} \times d(G,X)\right]}, \tag{A1}$$

where the predicted residuals $d(X,G)$ are given by

$$d(G,X) \equiv G\bar{h}(X) - \mathbb{L}[Gh(Z) \mid G,X] = G\bar{h}(X) - \theta_0 G - \sum_{j=1}^{J}\theta_j m_j(X), \tag{A2}$$

with coefficients $(\theta_0, ..., \theta_J)$ determined by

$$\mathbb{E}\left[G\left(G\bar{h}(X) - \theta_0 G - \sum_{j=1}^{J}\theta_j m_j(X)\right)\right] = 0, \tag{A3}$$

and

$$\mathbb{E}\left[m_j(X)\left(G\bar{h}(X) - \theta_0 G - \sum_{l=1}^{J}\theta_l m_l(X)\right)\right] = 0, \tag{A4}$$

for each $j = 1, ..., J$. The numerator of β can then be rewritten as

$$\begin{aligned}
\mathbb{E}\left[\Delta y \times d(G,X)\right] &= \mathbb{E}\left[d(G,X)\,\mathbb{E}[\Delta y \mid G,X]\right] \\
&= \mathbb{E}\left[d(G,X)\left(\mathbb{E}[\Delta \log Y(0) \mid G,X] + G\mathbb{1}[X \geq \bar{X}]\,\mathbb{E}[\log Y(1) - \log Y(0) \mid G,X]\right)\right], \\
&= \mathbb{E}[f^{y}(X)d(G,X)] + \mathbb{P}(G=1)\,\mathbb{E}[\mathbb{1}[X \geq \bar{Y}]d(G,X)[\log Y(1) - \log Y(0)] \mid G=1]
\end{aligned}$$

where the first equality follows from the law of iterated expectations and the fact that $d(G,X)$ is deterministic conditional on G and X, and the second equality follows from equation (5). The third equality follows from Assumption 1, equation (A3), and the law of total probability.

By Assumption 2 (exclusion), the last component of the expression can be rewritten as

$$\mathbb{E}[\mathbb{1}[X \geq \bar{Y}]d(G,X)[\log Y(1) - \log Y(0)] \mid G=1] = \mathbb{E}[\mathbb{1}[X \geq \bar{Y}]d(G,X)\Phi\zeta \mid G=1],$$

Similar arguments can be used to show that the denominator of β can be rewritten as

$$\mathbb{E}\left[\Delta\,\text{NTR}\times d(G,X)\right]=\mathbb{E}[f^{\text{NTR}}(X)d(G,X)]+\mathbb{P}(G=1)\,\mathbb{E}[\mathbb{1}[X\geq\bar{Y}]d(G,X)\Phi\mid G=1]$$

The result follows immediately. $\square$

Lemma 2. $\mathbb{E}[\bar{f}(X)d(G,X)]=0$ *for any function* $\bar{f}$ *if and only if* f *satisfies equation* (10).

Proof of Lemma 2. **Sufficiency:** To establish the sufficiency direction, suppose that f satisfies equation (10). Thus,

$$\mathbb{E}[d(G,X)\mid X]=\mathbb{E}[G\bar{h}(X)-\mathbb{L}[G,X]\mid X]=0 \tag{A5}$$

for each X. By the law of iterated expectations,

$$\mathbb{E}[\bar{f}(X)d(G,X)]=\mathbb{E}\left[\mathbb{E}[\bar{f}(X)d(G,X)\mid X]\right]=\mathbb{E}\left[\bar{f}(X)\,\mathbb{E}[d(G,X)\mid X]\right]=0.$$

Necessity: If f does not satisfy equation (10), then $\mathbb{E}[d(G,X)\mid X=x]\neq 0$ for at least one $x\in\mathcal{X}$. Let $D(x)\equiv\mathbb{E}[d(G,X)\mid X=x]$ and set

$$\bar{f}(x)\equiv\begin{cases}1 & \text{if } D(x)\geq 0,\\ -1 & \text{if } D(x)<0.\end{cases}$$

Using the law of iterated expectations, we obtain:

$$\begin{aligned}\mathbb{E}[\bar{f}(X)d(G,X)]&=\mathbb{P}(D(X)\geq 0)\,\mathbb{E}[\bar{f}(X)d(G,X)\mid D(X)\geq 0]+\mathbb{P}(D(X)<0)\,\mathbb{E}[\bar{f}(X)d(G,X)\mid D(X)<0],\\ &=\mathbb{P}(D(X)\geq 0)\,\mathbb{E}[d(G,X)\mid D(X)\geq 0]-\mathbb{P}(D(X)<0)\,\mathbb{E}[d(G,X)\mid D(X)<0],\end{aligned}$$

where the second equality follows from the definition of $\bar{f}$. By the law of iterated expectations:

$$\mathbb{E}[d(G,X)\mid D(X)\geq 0]=\mathbb{E}[\mathbb{E}[d(G,X)\mid X]\mid D(X)\geq 0]=\mathbb{E}[D(X)\mid D(X)\geq 0],$$

and similarly, $\mathbb{E}[d(G,X)\mid D(X)<0]=\mathbb{E}[D(X)\mid D(X)<0]$. Thus,

$$\mathbb{E}[\bar{f}(X)d(G,X)]=\mathbb{P}(D(X)>0)\,\mathbb{E}[D(X)\mid D(X)>0]-\mathbb{P}(D(X)<0)\,\mathbb{E}[D(X)\mid D(X)<0], \tag{A6}$$

where the strict inequalities in the first term are obtained using that $\mathbb{E}[D(X)\mid D(X)\geq 0]=\mathbb{P}(D(X)>0\mid D(X)\geq 0)\,\mathbb{E}[D(X)\mid D(X)>0]$. Since f does not satisfy equation (10), then $\mathbb{P}(D(X)\neq 0)=\mathbb{P}(D(X)>0)+\mathbb{P}(D(X)<0)>0$. By inspection of equation (A6), we then see that $\mathbb{E}[\bar{f}(X)d(G,X)]>0$. $\square$

Proof of Proposition 1. **Sufficiency:** To establish the sufficiency direction, suppose that f satisfies equation (10) and that $h(Z)$ takes two values. By Lemma 2, $\mathbb{E}[f^{y}(X)d(G,X)]=\mathbb{E}[f^{\text{NTR}}(X)d(G,X)]=0$ and by Lemma 1

$$\begin{aligned}\beta&=\frac{\mathbb{E}[\mathbb{1}[X\geq\bar{Y}]d(G,X)\Phi\zeta\mid G=1]}{\mathbb{E}[\mathbb{1}[X\geq\bar{Y}]d(G,X)\Phi\mid G=1]},\\ &=\frac{\sum_{x,\phi}\mathbb{1}[x\geq\bar{Y}]\,\mathbb{P}(X=x,\Phi=\phi\mid G=1)\,\mathbb{E}\left[d(G,X)\Phi\zeta\mid G=1,X=x,\Phi=\phi\right]}{\sum_{x,\phi}\mathbb{1}[x\geq\bar{Y}]\,\mathbb{P}(X=x,\Phi=\phi\mid G=1)\,\mathbb{E}\left[d(G,X)\Phi\mid G=1,X=x,\Phi=\phi\right]}\\ &=\frac{\sum_{x,\phi}\mathbb{1}[x\geq\bar{Y}]d(1,x)\,\mathbb{P}(X=x,\Phi=\phi\mid G=1)\phi\zeta(x,\phi)}{\sum_{x,\phi}\mathbb{1}[x\geq\bar{Y}]d(1,x)\,\mathbb{P}(X=x,\Phi=\phi\mid G=1)\phi},\end{aligned}$$

where the second equality uses the law of iterated expectations and the third follows from the definition of

$\zeta(x,\phi)$. By rearranging, we see that β recovers a weighted average of group-specific average ETIs:

$$\beta = \sum_{x,\phi} \underbrace{\frac{\mathbb{1}[x \geq \bar{Y}]d(1,x)\,\mathbb{P}(X=x,\Phi=\phi \mid G=1)\phi}{\sum_{x',\phi'} \mathbb{1}[x' \geq \bar{Y}]d(1,x')\,\mathbb{P}(X=x',\Phi=\phi' \mid G=1)\phi'}}_{\text{weights that sum to one}} \times \overbrace{\zeta(x,\phi)}^{\text{group-specific average ETIs}}. \tag{A7}$$

The weights clearly sum to one. Since probabilities are always weakly positive and assumption 3 ensures that $\phi \geq 0$, the weights are positive if $d(1,x)$ does not change sign with x.

By assumption, $h(Z)$ takes only two values. Since h is monotonic and strictly increasing at zero, we can write

$$h(Z) = h_0 + (h_1 - h_0)\mathbb{1}[Z > 0] = h_0 + (h_1 - h_0)\mathbb{1}[X \geq \bar{Y}].$$

for constants some constants $h_1 > h_0$. Combining this with the result from above that $\mathbb{E}[d(G,X) \mid X] = 0$, we get that

$$\sum_{j=1}^{J} \theta_j m_j(X) = \mathbb{E}[G \mid X]\left(h_0 + (h_1 - h_0)\mathbb{1}[X \geq \bar{Y}] - \theta_0\right)$$

Substituting this into the expression for $d(G,X)$ in equation (A2) and rewriting gives,

$$\begin{aligned} d(1,X) &= \left(h_0 + (h_1 - h_0)\mathbb{1}[X \geq \bar{Y}] - \theta_0\right)(1 - \mathbb{E}[G \mid X]) \\ &= \left(h_0 + (h_1 - h_0)\mathbb{1}[X \geq \bar{Y}] - \theta_0\right)\mathbb{P}(G = 0 \mid X). \end{aligned}$$

Substituting this into equation (A7), we obtain

$$\beta = \sum_{x,\phi} \frac{\mathbb{1}[x \geq \bar{Y}]\,\mathbb{P}(G=0 \mid X=x)\,\mathbb{P}(X=x,\Phi=\phi \mid G=1)\phi}{\sum_{x',\phi'} \mathbb{1}[x' \geq \bar{Y}]\,\mathbb{P}(G=0 \mid X=x')\,\mathbb{P}(X=x',\Phi=\phi' \mid G=1)\phi'} \times \zeta(x,\phi). \tag{A8}$$

which shows that the weights are positive and establishes that the ETI estimand is causal.

Necessity: To establish that $h(Z)$ taking only two values and f satisfying equation (10) are necessary conditions, we start by showing that if β is causal, then f satisfies (10). Suppose, for contradiction, that β is causal and f does not satisfy (10). By definition 1,

$$\beta = \sum_{x,\phi} \omega(x,\phi) \times \zeta(x,\phi),$$

for any collection of group-specific average ETIs. Consider the collection such that $\zeta(x,\phi) = \tilde{\zeta}$ for all x,ϕ and note that

$$\beta = \sum_{x,\phi} \omega(x,\phi) \times \zeta(x,\phi) = \tilde{\zeta} \sum_{x,\phi} \omega(x,\phi) = \tilde{\zeta},$$

since the weights sum to one. In the special case where

$$f^{\text{NTR}}(X) = \sum_{j=1}^{J} \theta'_j m_j(X),$$

for some vector $\theta' \in \mathbb{R}^J$, equation (A4) ensures that $\mathbb{E}[f^{\text{NTR}}(X)d(G,X)] = 0$. Thus, by Lemma 1,

$$\beta = \frac{\mathbb{E}[\mathbb{1}[X \geq \bar{Y}]d(G,X)\Phi\zeta \mid G=1]}{\mathbb{E}[\mathbb{1}[X \geq \bar{Y}]d(G,X)\Phi \mid G=1]} + \frac{\mathbb{E}[f^y(X)d(G,X)]}{\mathbb{E}[\Delta\,\text{NTR}\,d(G,X)]}.$$

By the law of iterated expectations

$$\begin{aligned}\mathbb{E}[\mathbb{1}[X \geq \bar{Y}]d(G,X)\Phi\zeta \mid G=1] &= \mathbb{E}[\mathbb{E}[\mathbb{1}[X \geq \bar{Y}]d(G,X)\Phi\zeta \mid G=1, X, \Phi] \mid G=1],\\ &= \mathbb{E}[\mathbb{1}[X \geq \bar{Y}]d(G,X)\Phi\,\mathbb{E}[\zeta \mid G=1, X, \Phi] \mid G=1],\\ &= \mathbb{E}[\mathbb{1}[X \geq \bar{Y}]d(G,X)\Phi\zeta(X,\Phi) \mid G=1],\\ &= \tilde{\zeta}\,\mathbb{E}[\mathbb{1}[X \geq \bar{Y}]d(G,X)\Phi \mid G=1],\end{aligned}$$

which gives

$$\beta = \tilde{\zeta} + \frac{\mathbb{E}[f^y(X)d(G,X)]}{\mathbb{E}[\Delta\,\text{NTR}\,d(G,X)]}.$$

Since β is causal, then $\mathbb{E}[f^y(X)d(G,X)] = 0$ for any arbitrary function f^y. By Lemma 2, this implies that f satisfies the condition in (10), leading to a contradiction. We conclude that if β is causal, then f satisfies 10.

To show that $h(Z)$ taking only two values is necessary, we note that when f satisfies (10), then the ETI estimand is given by equation (A7), and the sign of the weights is determined by $d(1,X)$:

$$d(1,X) = (\bar{h}(X) - \theta_0)\,\mathbb{P}(G=0 \mid X).$$

To derive an expression for θ_0, we use that since $\mathbb{E}[d(G,X) \mid X] = 0$,

$$\sum_{j=1}^{J} \theta_j\,\mathbb{E}[m_j(X) \mid G=1] = \mathbb{E}\left[\mathbb{E}[G \mid X]\left(\bar{h}(X) - \theta_0\right) \mid G=1\right],$$

and by equation (A3),

$$\sum_{j=1}^{J} \theta_j\,\mathbb{E}[m_j(X) \mid G=1] = \mathbb{E}[\bar{h}(X) \mid G=1] - \theta_0.$$

Solving for θ_0 gives:

$$\theta_0 = \frac{\mathbb{E}[\mathbb{P}(G=0 \mid X)\bar{h}(X) \mid G=1]}{\mathbb{E}[\mathbb{P}(G=0 \mid X) \mid G=1]},$$

meaning that the sign of $d(1,X)$ is determined by the sign of

$$\bar{h}(X) - \frac{\mathbb{E}[\mathbb{P}(G=0 \mid X)\bar{h}(X) \mid G=1]}{\mathbb{E}[\mathbb{P}(G=0 \mid X) \mid G=1]}.$$

To see that this can generally take on positive and negative values, consider a distribution of initial income X with support $\{x_1, x_2, x_3\}$, where $x_1 < \bar{Y} \leq x_2, x_3$ and $\bar{h}(x_1) < \bar{h}(x_2) < \bar{h}(x_3)$ and G is independent of X. In this case, $d(1,x)$ simplifies to

$$d(1,x) = \bar{h}(x) - \left(\mathbb{P}(X=x_1)\bar{h}(x_1) + \mathbb{P}(X=x_2)\bar{h}(x_2) + (1 - \mathbb{P}(X=x_1) - \mathbb{P}(X=x_2))\bar{h}(x_3)\right).$$

Thus, if there exits probabilities $\mathbb{P}(X = x_1)$ and $\mathbb{P}(X = x_2)$ such that

$$\bar{h}(x_2) < \mathbb{P}(X = x_1)\bar{h}(x_1) + \mathbb{P}(X = x_2)\bar{h}(x_2) + (1 - \mathbb{P}(X = x_1) - \mathbb{P}(X = x_2))\bar{h}(x_3), \tag{A9}$$

$$\bar{h}(x_3) > \mathbb{P}(X = x_1)\bar{h}(x_1) + \mathbb{P}(X = x_2)\bar{h}(x_2) + (1 - \mathbb{P}(X = x_1) - \mathbb{P}(X = x_2))\bar{h}(x_3), \tag{A10}$$

with $\mathbb{P}(X = x_1) \in (0, 1)$, $\mathbb{P}(X = x_2) \in (0, 1)$ and $\mathbb{P}(X = x_1) + \mathbb{P}(X = x_2) \in (0, 1)$, we can conclude that positive weights are not guaranteed.

Since $\bar{h}(x_1) < \bar{h}(x_2) < \bar{h}(x_3)$, the inequality in (A10) is always satisfied. The inequality in (A9) is satisfied whenever

$$0 < \mathbb{P}(X = x_1) < (1 - \mathbb{P}(X = x_2))\frac{\bar{h}(x_3) - \bar{h}(x_2)}{\bar{h}(x_3) - \bar{h}(x_1)} \qquad \text{and} \qquad \mathbb{P}(X = x_2) \in (0, 1),$$

where the region is nonempty because $[\bar{h}(x_3) - \bar{h}(x_2)]/[\bar{h}(x_3) - \bar{h}(x_1)] > 0$.

In conclusion, negative weights are not ruled out unless h is binary. Lastly, by Lemma 3, the weights are unique, so that there exists no other set of weights that are positive and sum to one, meaning that β is not causal. □

Lemma 3. *If $\beta = \sum_{x,\phi} \omega(x, \phi) \times \zeta(x, \phi)$ for all $\bar{\zeta}$, then the weights are unique.*

Proof of Lemma 3. To prove that if $\beta = \sum_{x,\phi} \omega(x, \phi) \times \zeta(x, \phi)$ and $\beta = \sum_{x,\phi} \omega'(x, \phi) \times \zeta(x, \phi)$ for any $\bar{\zeta}$ then $\omega(x, \phi) = \omega'(x, \phi)$, suppose for contradiction that $\omega(x, \phi) \neq \omega'(x, \phi)$ for at least one $(x, \phi) = (x^*, \phi^*)$. By hypothesis, we can write

$$\sum_{x,\phi}(\omega(x, \phi) - \omega'(x, \phi))\zeta(x, \phi) = 0. \tag{A11}$$

By assumption, there exists at least one group (x^*, ϕ^*) with $\omega(x^*, \phi^*) \neq \omega'(x^*, \phi^*)$. Consider the $\bar{\zeta}$ characterized by

$$\zeta(x, \phi) = \begin{cases} 1 & \text{if } (x, \phi) = (x^*, \phi^*), \\ 0 & \text{otherwise.} \end{cases}$$

Then,

$$\sum_{x,\phi}(\omega(x, \phi) - \omega'(x, \phi))\zeta(x, \phi) = \omega(x^*, \phi^*) - \omega'(x^*, \phi^*) \neq 0,$$

contradicting (A11). We conclude that $\omega(x, \phi) = \omega'(x, \phi)$ for all x, ϕ. □

Proof of Corollary 1. **Sufficiency:** Since f satisfies (10), Lemma 2 implies that $\mathbb{E}[f^y(X)d(G, X)] = \mathbb{E}[f^{\text{NTR}}(X)d(G, X)] = 0$. By Lemma 1,

$$\beta = \frac{\mathbb{E}[\mathbb{1}[X \geq \bar{Y}]d(G, X)\Phi\zeta \mid G = 1]}{\mathbb{E}[\mathbb{1}[X \geq \bar{Y}]d(G, X)\Phi \mid G = 1]} = \zeta \times \frac{\mathbb{E}[\mathbb{1}[X \geq \bar{Y}]d(G, X)\Phi \mid G = 1]}{\mathbb{E}[\mathbb{1}[X \geq \bar{Y}]d(G, X)\Phi \mid G = 1]} = \zeta.$$

Necessity: The first part of the necessity proof of Proposition 1 demonstrated that if $\zeta(x, \phi) = \tilde{\zeta}$ for all x, ϕ, then $\beta = \tilde{\zeta}$ only if f satisfies (10). When ζ is homogenous across individuals, then $\zeta(x, \phi) = \zeta$ for all x, ϕ, and $\beta = \zeta$ only if f satisfies (10). □

Proof of Corollary 2. To derive the expression for the weights $\omega(x, \phi)$, start by noting that

$$\mathbb{P}(X, \Phi \mid G = 1) = \mathbb{P}(\Phi \mid G = 1, X)\,\mathbb{P}(X \mid G = 1) = \mathbb{P}(\Phi \mid G = 1, X)\frac{\mathbb{P}(G = 1 \mid X)\,\mathbb{P}(X = x)}{\mathbb{P}(G = 1)}$$

Substituting this into equation (A8) gives

$$\beta = \sum_{x,\phi} \frac{\mathbb{1}[x \geq \bar{Y}]\,\mathbb{P}(G=0 \mid X=x)\,\mathbb{P}(G=1 \mid X=x)\,\mathbb{P}(X=x)\phi}{\sum_{x',\phi'} \mathbb{1}[x' \geq \bar{Y}]\,\mathbb{P}(G=0 \mid X=x')\,\mathbb{P}(G=1 \mid X=x')\,\mathbb{P}(X=x')\phi'} \times \zeta(x,\phi).$$

The result follows by using that

$$\mathbb{P}(G=0 \mid X=x)\,\mathbb{P}(G=1 \mid X=x) = Var(G \mid X=x),$$

since G is binary. □

The following Lemma will be useful in proving the results that involve the local ETI estimand. It shows how $\beta(x)$ is related to ζ, Φ and ν under Assumption 1.

Lemma 4. *Suppose Assumption 1 holds. Then,*

$$\beta(x) = \sum_{\phi} \frac{\mathbb{P}(\Phi=\phi \mid G=1, X=x) \times \phi}{\sum_{\phi'} \mathbb{P}(\Phi=\phi' \mid G=1, X=x) \times \phi'} \times \zeta(x,\phi) + \frac{\mathbb{E}[\nu(1)-\nu(0) \mid G=1, X=x]}{\mathbb{E}[\Delta\,\text{NTR} \mid G=1, X=x] - (\lambda_0^{\text{NTR}} + f^0(x;\lambda^{\text{NTR}}))}.$$

Proof of Lemma 4. Equation (5) implies that

$$\mathbb{E}[\Delta y \mid G, X] = \begin{cases} \mathbb{E}[\Delta \log Y(0) + \log Y(1) - \log Y(0) \mid G, X] & \text{if } G=1 \text{ and } X \geq \bar{Y}, \\ \mathbb{E}[\Delta \log Y(0) \mid G, X] & \text{otherwise.} \end{cases}$$

This means that

$$\mathbb{E}[\Delta y \mid G, X, GZ=0] = \mathbb{E}[\Delta \log Y(0) \mid G, X] = \kappa^y G + f^y(X).$$

where the second equality uses assumption 1. Because f^0 is saturated, $\lambda_0^y = \kappa^y$ and $f^0(X;\lambda^y) = f^y(X)$, then,

$$\mathbb{E}[\Delta \log Y(0) \mid G, X] = \lambda_0^y G + f^0(X;\lambda^y).$$

Thus, for any $x \geq \bar{Y}$, the numerator of $\beta(x)$ equals,

$$\begin{aligned} &\mathbb{E}[\Delta y \mid G=1, X=x] - \left(\lambda_0^y + f^0(x;\lambda^y)\right), \\ &= \mathbb{E}[\Delta \log Y(0) + \log Y(1) - \log Y(0) \mid G=1, X=x] - \mathbb{E}[\Delta \log Y(0) \mid G=1, X=x], \\ &= \mathbb{E}[\Phi \times \zeta \mid G=1, X=x] + \mathbb{E}[\nu(1)-\nu(0) \mid G=1, X=x], \\ &= \sum_{\phi} \mathbb{P}(\Phi=\phi \mid G=1, X=x) \times \phi \times \zeta(x,\phi) + \mathbb{E}[\nu(1)-\nu(0) \mid G=1, X=x], \end{aligned}$$

where the second equality follows from Assumption 1, and the third follows from the law of iterated expectations. Analogous arguments show that the denominator of the $\beta(x)$ equals,

$$\mathbb{E}[\Delta\,\text{NTR} \mid G=1, X=x] - \left(\lambda_0^{\text{NTR}} + f^0(X;\lambda^{\text{NTR}})\right) = \sum_{\phi} \mathbb{P}(\Phi=\phi \mid G=1, X=x) \times \phi.$$

The lemma follows by combining these expressions for the numerator and denominator of $\beta(x)$ and rearranging. □

Proof of Proposition 2. Under assumption 2, Lemma 4 directly implies the expression for $\beta(x)$ in the proposition. The weights clearly sum to one, and by Assumption 3, $\phi \geq 0$, implying that the weights are positive. □

A.2 Proofs of results from Section 4

Proof of Proposition 3. By equation (18), the potential earnings function is related to the uncompensated labor supply function through

$$\log Y^u(\tau, R(d) + B) = \zeta \log(1 - \tau) + \nu(d). \tag{A12}$$

Implicit differentiation of equation (A12) with respect to τ gives,

$$\frac{-\zeta}{1 - \tau} = \frac{1}{Y^u} \frac{\partial Y^u}{\partial \tau}.$$

Rearranging and using that $\partial Y^u / \partial(1 - \tau) = -\partial Y^u / \partial \tau$, we obtain

$$\zeta = \frac{\partial Y^u}{\partial 1 - \tau} \frac{1 - \tau}{Y^u} = \varepsilon^u,$$

with ε^u defined as in equation (16). Thus,

$$\log Y^u(\tau, R(d) + B) = \varepsilon^u \log(1 - \tau) + \nu(d), \tag{A13}$$

implying that

$$\nu(1) - \nu(0) = \int_{R(0)}^{R(1)} \frac{\partial \log Y^u(\tau, z + B)}{\partial R} dz = \int_{R(0)}^{R(1)} \frac{\eta(\tau, z + B)}{(1 - \tau) Y^u(\tau, z + B)} dz. \tag{A14}$$

Since income effects η are non-positive and $R(0) \neq R(1)$ for at least some individuals, this clearly shows that $\nu(0) = \nu(1)$ if and only if income effects are always zero. Moreover, if $\eta = 0$, then $\varepsilon^u = \varepsilon^c$ by the Slutsky equation.

□

Proof of Proposition 4. Equations (A13) and (A14) can be derived using the same arguments as in the first part of the proof of Proposition 3. By assumption, individuals do not switch brackets because of the tax reform, meaning that $\tau(d) = T'_d(Y(0)) = T'_d(Y(1))$ and

$$R(d) = T'_d(Y(0))Y(0) - T_d(Y(0)) = T'_d(Y(1))Y(1) - T_d(Y(1)),$$

for $d = 0, 1$. Thus, under no bracket switching, a tax reform that sets log net-of-marginal tax rates to equal NTR and tax liability equal to T for incomes Y gives the following virtual income at Y:

$$\tilde{R}(\text{NTR}, T) = [1 - \exp(\text{NTR})] \, Y - T,$$

meaning that violation of the exclusion restriction can be expressed as

$$\nu(1) - \nu(0) = \int_{\tilde{R}(\text{NTR}(0), T_0(Y))}^{R(1)} \frac{\eta(\tau(1), z + B)}{(1 - \tau(1)) Y^u(\tau(1), z + B)} dz.$$

Taking a first-order approximation around $\text{NTR}(1), T_1(Y(1))$ gives:

$$\begin{aligned}
\nu(1)-\nu(0) &\approx -\left(\frac{\eta}{(1-\tau)Y}\right)\left(\frac{\partial \tilde{R}}{\partial\,\text{NTR}}(\text{NTR}(0)-\text{NTR}(1)) + \frac{\partial \tilde{R}}{\partial T}(T_0(Y(1)) - T_1(Y(1)))\right), \\
&= \left(\frac{\eta}{(1-\tau)Y}\right)\Big(-\exp(\text{NTR})Y\Phi - (T_1(Y)-T_0(Y))\Big), \\
&= -\eta\left(\Phi + \frac{T_1(Y)-T_0(Y)}{(1-\tau)Y}\right)
\end{aligned}$$

where we have used that the observed post-reform τ, Y and η are equal to their potential counterparts evaluated under T_1. By Lemma 4,

$$\beta(x) = \varepsilon^u(x) - \eta(x) - \eta(x)\frac{\mathbb{E}\left[\frac{T_1(Y)-T_0(Y)}{(1-T_1'(Y))Y} \mid G=1, X=x\right]}{\mathbb{E}[\Delta\,\text{NTR} \mid G=1, X=x] - (\lambda_0^{\text{NTR}} + f^0(X;\lambda^{\text{NTR}}))}, \tag{A15}$$

where we have used that ε^u and η are constant conditional on X. Equation (20) follows from rearranging. Equation (19) follows from rearranging and using that $\varepsilon^u(x) - \eta(x) = \varepsilon^c(x)$ by the Slutsky equation.

□

B Appendix figures and tables

Table A1: Summary statistics for ETI sample

	Full sample	Estimation sample	
		Reform cohorts	Placebo cohorts
		Demographics	
Age	40.70	40.86	40.66
Education	12.14	12.32	12.29
Male	57.20	62.80	64.80
Married	61.60	57.70	61.30
		Income and taxes	
Taxable income X	443	470	468
Income taxes $T(X)$	130	132	139
Marginal tax rate $T'(X)$	0.414	0.425	0.430
Observations	5,702,759	2,286,662	1,677,302
Individuals	1,349,358	955,572	796,533

Notes: This table reports summary statistics for the sample used in the estimation of the elasticity of taxable income (ETI). Monetary values are consumer-price-index adjusted and reported in 1,000 2018 NOKs (8.13 NOK/USD), and binary outcomes are reported in percent. Observations from 2002-2004 comprise the placebo cohorts, while observations from 2005-2007 comprise the reform cohorts.

Table A2: Robustness to employment responses

	Employment response	ETI	
		Lower bound	Upper bound
Estimate	0.0006	0.1814	0.2827
SE	0.0001	0.0130	0.0134

[a] This table presents employment response estimates to the reform, and bounds on the ETI. Employment responses are obtained from estimating the reduced form of the regression model in equations (3) and (4) on a sample that includes the unemployed, where the outcome is a binary variable equal to one if the individual works and zero otherwise. The specification employs a binary instrument and controls for initial income X using percentile dummies. Standard errors for the ETI bounds are computed via bootstrapping with 500 replications. See Appendix C for details and the derivation of the ETI bounds.

Table A3: Robustness

Specification		(i)	(ii)	(iii)
First stage	Coef.	0.0517	0.0507	0.0526
	SE	0.0002	0.0002	0.0002
Reduced form	Coef.	0.0107	0.0090	0.0122
	SE	0.0006	0.0007	0.0006
ETI	Coef.	0.2079	0.1784	0.2325
	SE	0.0122	0.0133	0.0121

[a] This table presents ETI estimates for different samples and outcomes. Column (i) excludes self-employed and owners of closely held corporations, column (ii) uses total market income as the outcome, and column (iii) weights by initial income. All specifications set $h(Z) = \mathbb{1}[Z > 0]$ and control for initial income using a full set of percentile dummies in X.

(a) Employment effects. (b) Bounds on local ETIs

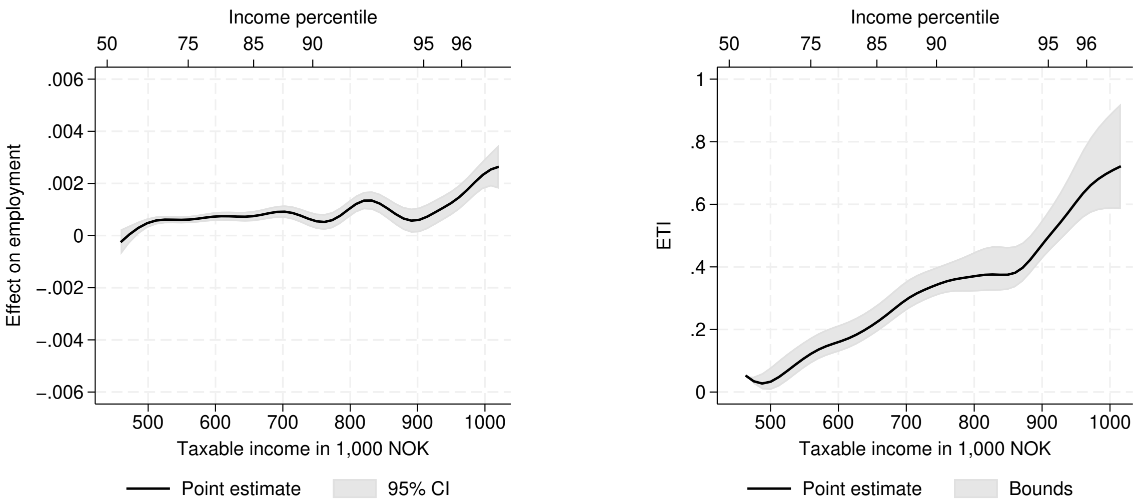


Figure A1: Robustness to employment responses

Notes: The figure plots estimates of the local employment effects and bounds on the ETI across the income distribution. Panel (a) plots the local employment effects obtained by first estimating the counterfactual change in employment using 100 percentile bins of X, then estimating the numerator of equation (11) using a local regression, with the change in employment replacing the change in log incomes. Panel (b) plots the bounds on the local ETIs as explained in Appendix C. 95 percent confidence intervals are shown, with standard errors obtained by bootstrapping the entire estimation procedure using 500 replications.

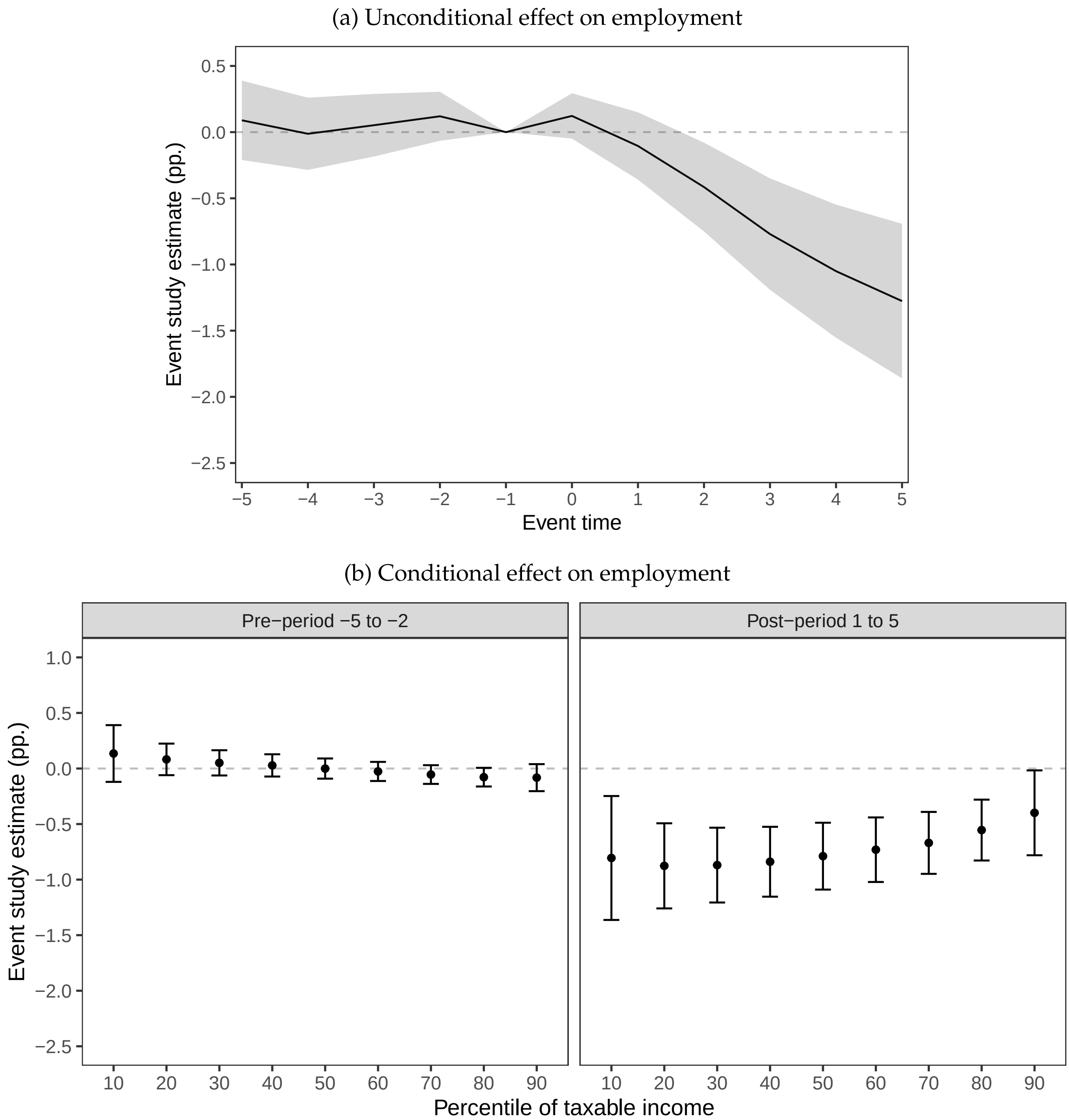


Figure A2: Employment effects of winning

Notes: This figure presents estimates of the effect of winning the lottery on winners' employment. Panel (a) presents unconditional estimates of the effect of winning the lottery on winners' employment for each event time t. The estimates correspond to the sample analogue of equation (25) using employment as an outcome instead of earnings, evaluated without conditioning on X, controlling for age, and averaged across cohorts using cohort-size weights. The estimates in Panel (b) correspond to the sample analogues of equation (25) using employment as an outcome instead of earnings, evaluated at different values of X. We use a Gaussian kernel with a bandwidth of 100,000 NOK, where the evaluation points $\{x\}$ correspond to the cohort-specific deciles of pre-win taxable income. For each decile, we compute cohort-size weighted averages across cohorts and report equally weighted averages across event times $t = -5$ to -2 and $t = +1$ to $+5$. 95 percent confidence intervals are shown, with standard errors clustered at the winner level. Throughout, we use $g - 1$ as the omitted event time.

Table A4: Summary statistics for lottery and ETI estimation samples

	Lottery sample		ETI sample	Stand. mean diff.	
	Original	Re-weighted		Before	After
Age	43.97	40.83	40.78	0.36	0.01
Education	11.86	12.31	12.30	-0.17	0.00
Male	58.26	63.91	63.62	-0.11	0.01
Married	58.43	57.94	59.20	-0.02	-0.03
Taxable income X	455	469	469	-0.07	0.00
Income taxes $T(X)$	131	132	134	-0.04	-0.03
Marginal tax rate $T'(X)$	0.40	0.43	0.43	-0.29	0.00
Observations	1,182,715	366,961	4,011,294		
Individuals	18,413	11,652	1,126,813		
ESS		298,438			

Notes: This table reports summary statistics for the lottery sample and the ETI estimation sample in the reference period, with the lottery sample shown both before and after reweighting. Monetary variables are consumer-price-index adjusted and reported in 1,000 2018 NOK, and binary variables are reported in percent. The ETI observations correspond to the combined placebo and reform cohorts used in Table A1. The original and reweighted lottery samples are based on different winner cohorts, since reweighting is implemented for cohorts aligned with the ETI design. Weights are based on propensity scores from a logit model with dummies for year, age, gender, marital status, schooling, and taxable-income bins. For the reweighted lottery sample, we apply the same sample restrictions as in the ETI estimation sample. Standardized mean differences before and after reweighting are computed as differences in means divided by the square root of the average of the two sample variances. ESS denotes the effective sample size of the reweighted lottery sample.

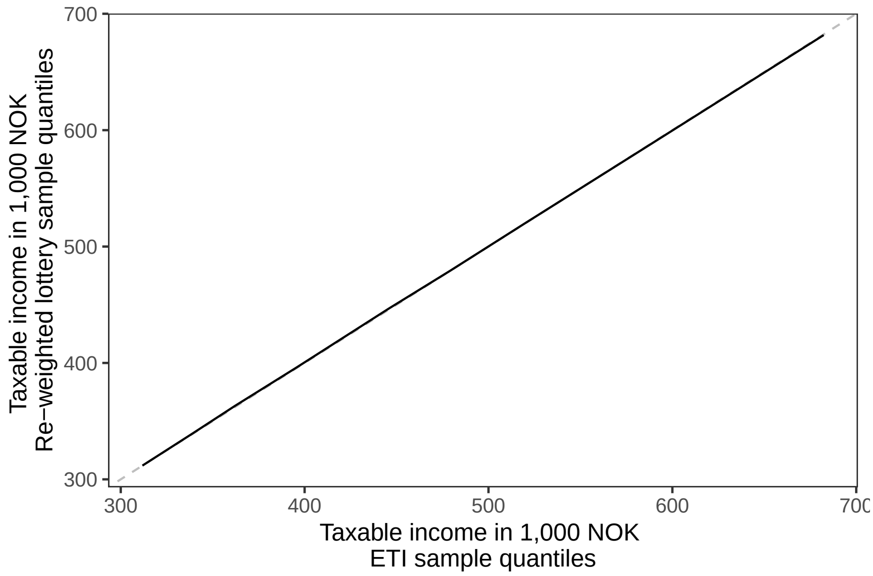


Figure A3: Distribution of taxable income in the ETI and reweighted lottery samples

Notes: This figure plots quantiles of taxable (initial) income in the ETI estimation sample against those in the reweighted lottery sample, measured in each sample's respective reference period. The dashed 45-degree line indicates equality of the two quantiles.

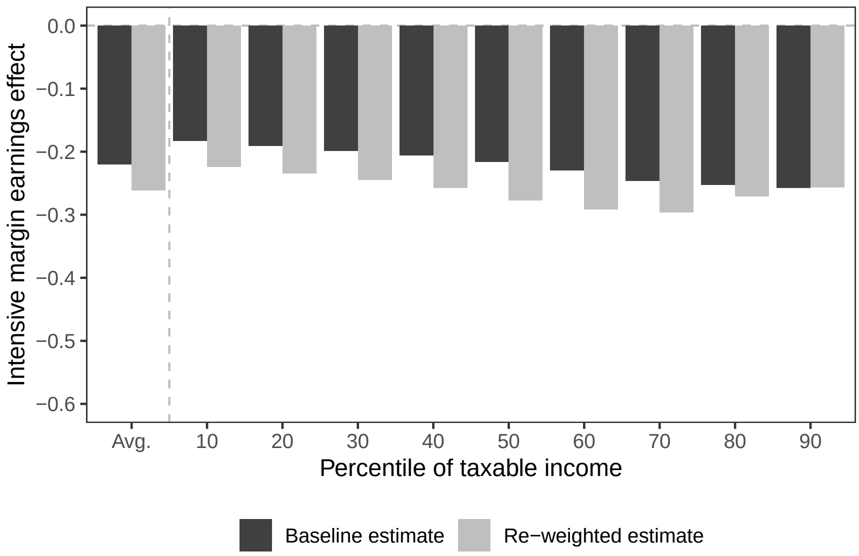


Figure A4: Robustness of intensive-margin income effects to reweighting

Notes: This figure plots the intensive-margin income effects conditional on taxable income in the original lottery sample and the reweighted lottery sample. Taxable income is measured in the year prior to winning the lottery. The effects are estimated as in Figure 7, panel (b), using the same post-treatment window $+1$ to $+5$.

C Bounds on ETIs when allowing for employment responses

In the main specifications, we condition on positive earnings after the reform. If the tax reform induces employment responses, this could mean that the sample used to estimate ETIs is endogenous to the tax reform. This appendix presents a method for computing bounds on the ETI estimand that accounts for this.

We consider the same potential outcomes model as in Section 2 and introduce the following exhaustive and mutually exclusive typology of extensive-margin response types defined from

$$S \equiv \begin{cases} \text{AW} & \text{if } Y(0) > 0 \text{ and } Y(1) > 0, \\ \text{EN} & \text{if } Y(0) = 0 \text{ and } Y(1) > 0, \\ \text{EX} & \text{if } Y(0) > 0 \text{ and } Y(1) = 0, \\ \text{NW} & \text{if } Y(0) = 0 \text{ and } Y(1) = 0. \end{cases}$$

We refer to individuals with $S = \text{AW}$ as always-workers, individuals with $S = \text{EN}$ as entry-workers, individuals with $S = \text{EX}$ as exit-workers, and individuals with $S = \text{NW}$ as never-workers. Because the tax reform increased incentives to work, we rule out the existence of exit-workers by assuming that $\mathbb{P}(S = \text{EX}) = 0$.

C.1 Bounds on local ETI estimands

If S was observable in the data, one could simply condition on $S = \text{AW}$ and estimate the local ETIs as before. In this case, the local ETI estimand would be given by

$$\beta^{\text{AW}}(x) \equiv \frac{\mathbb{E}[\Delta y \mid G = 1, X = x, S = \text{AW}] - \left(\lambda_0^y + f^0(x; \lambda^y)\right)}{\mathbb{E}[\Delta\, \text{NTR} \mid G = 1, X = x, S = \text{AW}] - \left(\lambda_0^{\text{NTR}} + f^0(x; \lambda^{\text{NTR}})\right)},$$

where

$$(\lambda_0^w, \lambda^w) \equiv \underset{\lambda_0, \lambda}{\arg\min} \quad \mathbb{E}\left[\left(\Delta w - \lambda_0 G - f^0(x; \lambda^w)\right)^2 \mid GZ = 0, S = \text{AW}\right],$$

for $w = y, \text{NTR}$. By imposing Assumptions 1 - 3 conditional on $S = \text{AW}$, one can then apply Proposition 2 to get that $\beta^{\text{AW}}(x)$ would recover a positively weighted ETI conditional on $S = \text{AW}$.

Because S is not observable in the data, this simple approach is not feasible. However, it is possible to make some progress by noting that individuals with positive earnings who do not experience the tax reform (i.e., have $GZ = 0$) are always-workers. To see why, recall that $\mathbb{P}(S = \text{EX}) = 0$, which means that $Y(0) > 0$ implies that $S = \text{AW}$. As a consequence, $\hat{\lambda}^y$ and $\hat{\lambda}^{\text{NTR}}$ as defined in equation (11) are already estimated on the sample of always-workers.

To learn about $\mathbb{E}[\Delta y \mid G = 1, X = x, S = \text{AW}]$ and $\mathbb{E}[\Delta\, \text{NTR} \mid G = 1, X = x, S = \text{AW}]$, we rely on arguments from Horowitz and Manski (1995) and Lee (2005). Letting $E \equiv \mathbb{1}[Y > 0]$, it is useful to observe that the distribution of earnings growth conditional on positive earnings is a mixture of the distribution of earnings growth for always- and entry-workers:

$$\begin{aligned}\underbrace{\mathbb{P}(\Delta y \le z \mid G = 1, X = x, E = 1)}_{\text{observed}} &= \overbrace{(1 - p(x))}^{\text{estimated}} \underbrace{\mathbb{P}(\Delta y \le z \mid G = 1, X = x, S = \text{AW})}_{\text{unknown}} \\ &\quad + \overbrace{p(x)}^{\text{estimated}} \underbrace{\mathbb{P}(\Delta y \le z \mid G = 1, X = x, S = \text{EN})}_{\text{unknown}}\end{aligned}$$

where $p(x) \equiv \mathbb{P}(S = \text{EN} \mid G = 1, X = x, Y(1) > 0)$. The LHS is directly observed, $p(x)$ can be estimated using the same design as we have already used to estimate the ETI. Specifically, the local employment effects estimated in Section 3.5 can be used to recover $p(x)$ through

$$p(x) = \frac{\overbrace{\mathbb{P}(S = \text{EN} \mid G = 1, X = x)}^{\text{estimated in Section 3}}}{\underbrace{\mathbb{P}(E = 1 \mid G = 1, X = x)}_{\text{observed}}}. \tag{A16}$$

Thus, the arguments from Horowitz and Manski (1995) and Lee (2005) applies and we obtain the following bounds for $w = y, \text{NTR}$:

$$LB^w(x) \leq \mathbb{E}[\Delta w \mid G = 1, X = x, S = \text{AW}] \leq UB^w(x),$$

where

$$LB^w(x) \equiv \mathbb{E}[\Delta w \mid G = 1, X = x, S = \text{AW}, \Delta w \leq q^w(1 - p(x); x)],$$
$$UB^w(x) \equiv \mathbb{E}[\Delta w \mid G = 1, X = x, S = \text{AW}, \Delta w > q^w(p(x); x)],$$

and q^w is the quantile function defined by $\mathbb{P}(\Delta w \leq q^w(p; x) \mid G = 1, X = x, E = 1) = p$.

Using these results, $\beta^{\text{AW}}(x)$, is bounded by

$$\frac{LB^y(x) - \left(\lambda_0^y + f^0(x; \lambda^y)\right)}{UB^{\text{NTR}}(x) - \left(\lambda_0^{\text{NTR}} + f^0(x; \lambda^{\text{NTR}})\right)} \leq \beta^{\text{AW}}(x) \leq \frac{UB^y(x) - \left(\lambda_0^y + f^0(x; \lambda^y)\right)}{LB^{\text{NTR}}(x) - \left(\lambda_0^{\text{NTR}} + f^0(x; \lambda^{\text{NTR}})\right)},$$

where λ_0^w and λ^w for $w = y, \text{NTR}$ are defined as equation (11), and we have used that $LB^{\text{NTR}}(x) > \lambda_0^{\text{NTR}} + f^0(x; \lambda^{\text{NTR}})$ in the data.

C.2 Bounds on the ETI estimand

To see how bounds on the ETI estimand for always-workers β^{AW} can be obtained, it is useful to observe that, by Corollary 2, β^{AW} is related to $\beta^{\text{AW}}(x)$ through

$$\beta^{\text{AW}} = \sum_x \frac{\mathbb{1}[x \geq \bar{Y}]\, \mathbb{P}(X = x) Var(G \mid X = x) FS^{\text{AW}}(x)}{\sum_{x'} \mathbb{1}[x' \geq \bar{Y}]\, \mathbb{P}(X = x') Var(G \mid X = x') FS^{\text{AW}}(x')} \beta^{\text{AW}}(x),$$
$$= \sum_x \frac{\mathbb{1}[x \geq \bar{Y}]\, \mathbb{P}(X = x) Var(G \mid X = x) RF^{\text{AW}}(x)}{\sum_{x'} \mathbb{1}[x' \geq \bar{Y}]\, \mathbb{P}(X = x') Var(G \mid X = x') FS^{\text{AW}}(x')}$$

with

$$RF^{\text{AW}}(x) \equiv \mathbb{E}[\Delta y \mid G = 1, X = x', S = \text{AW}] - \left(\lambda_0^y + f^0(x'; \lambda^y)\right),$$
$$FS^{\text{AW}}(x) \equiv \mathbb{E}[\Delta\, \text{NTR} \mid G = 1, X = x', S = \text{AW}] - \left(\lambda_0^{\text{NTR}} + f^0(x'; \lambda^{\text{NTR}})\right).$$

Thus, β^{AW} must satisfy

$$\beta_{\text{LB}} \leq \beta^{\text{AW}} \leq \beta_{\text{UB}}$$

with

$$\beta_{\text{LB}} \equiv \frac{\sum_x \mathbb{1}[x \geq \bar{Y}]\, \mathbb{P}(X = x) Var(G \mid X = x) \left[LB^y(x) - \left(\lambda_0^y + f^0(x; \lambda^y)\right)\right]}{\sum_{x'} \mathbb{1}[x' \geq \bar{Y}]\, \mathbb{P}(X = x') Var(G \mid X = x') \left[UB^{\text{NTR}}(x) - (\lambda_0^{\text{NTR}} + f^0(x'; \lambda^{\text{NTR}}))\right]},$$
$$\beta_{\text{UB}} \equiv \frac{\sum_x \mathbb{1}[x \geq \bar{Y}]\, \mathbb{P}(X = x) Var(G \mid X = x) \left[UB^y(x) - \left(\lambda_0^y + f^0(x; \lambda^y)\right)\right]}{\sum_{x'} \mathbb{1}[x' \geq \bar{Y}]\, \mathbb{P}(X = x') Var(G \mid X = x') \left[LB^{\text{NTR}}(x) - (\lambda_0^{\text{NTR}} + f^0(x'; \lambda^{\text{NTR}}))\right]},$$

where λ_0^w and λ^w for $w = y, \text{NTR}$ are defined as equation (11), and we have used that $LB^{\text{NTR}}(x) > \lambda_0^{\text{NTR}} + f^0(x; \lambda^{\text{NTR}})$ in the data.

D Formal identification results for the lottery design

This appendix derives the formal identification results underlying our empirical strategy for the lottery design. We first map potential outcomes to observed data, show that the reduced-form estimand identifies the causal earnings effect under a conditional parallel-trends assumption, and then establish the decomposition of the total earnings response into intensive- and extensive-margin components.

Mapping potential outcomes to observed data. Let $W_t(d)$ denote the potential outcome W in calendar year t if the individual has already won ($d = 1$) or not won ($d = 0$) for $W \in \{Y, B\}$, where Y and B denote earnings and unearned income, respectively. Potential outcomes map to the observed data through

$$W_t = \mathbb{1}[G \leq t]\big(W_t(1) - W_t(0)\big) + W_t(0) \tag{A17}$$

Earnings, employment, and unearned income effects of winning the lottery. We start by stating the parallel trends assumption we rely on for obtaining the effect of winning the lottery on earnings and unearned income. It says that conditional on pre-win earnings $Y_{g-1} = x$, average earnings and unearned income for winners and later-winners would have evolved in parallel:

$$\begin{aligned} &\mathbb{E}[W_{g+t}(0) - W_{g-1}(0) \mid Y_{g-1} = x,\, G = g] \\ &\qquad\qquad = \mathbb{E}\left[W_{g+t}(0) - W_{g-1}(0) \mid Y_{g-1} = x,\, G > g + t\right], \end{aligned} \tag{A18}$$

for all $t \geq 0$ and $W \in \{Y, B\}$. Under this assumption, the reduced form estimand $\text{RF}_{g,t}(x)$ ($\text{FS}_{g,t}(x)$) recovers the average earnings (unearned income) effect of winning the lottery for cohort g. To see why, note that

$$\begin{aligned} \text{RF}_{g,t}(x) &\equiv \mathbb{E}[Y_{g+t} - Y_{g-1} \mid Y_{g-1} = x, G = g] - \mathbb{E}[Y_{g+t} - Y_{g-1} \mid Y_{g-1} = x, G > g + t] \\ &= \mathbb{E}[Y_{g+t}(1) - Y_{g-1}(0) \mid Y_{g-1} = x, G = g] - \mathbb{E}[Y_{g+t}(0) - Y_{g-1}(0) \mid Y_{g-1} = x, G > g + t] \\ &= \mathbb{E}[Y_{g+t}(1) - Y_{g+t}(0) \mid Y_{g-1} = x, G = g] \\ &\quad + \Big(\mathbb{E}[Y_{g+t}(0) - Y_{g-1}(0) \mid Y_{g-1} = x, G = g] - \mathbb{E}[Y_{g+t}(0) - Y_{g-1}(0) \mid Y_{g-1} = x, G > g + t]\Big), \end{aligned}$$

for all $t \geq 0$. By the parallel-trends assumption (A18), the term in parentheses equals zero, and hence

$$\text{RF}_{g,t}(x) = \mathbb{E}[Y_{g+t}(1) - Y_{g+t}(0) \mid Y_{g-1} = x,\, G = g]\,.$$

Analogous arguments apply to unearned income and employment.

Aggregating across cohorts and event times. We aggregate the cohort and event-time specific reduced form and first stage estimands according to

$$\text{RF}(x) \equiv \sum_{g \in \mathcal{G}} \mathbb{P}(G = g \mid Y_{g-1} = x)\, \frac{1}{5} \sum_{t=1}^{5} \text{RF}_{g,t}(x), \tag{A19}$$

$$\text{FS}(x) \equiv \sum_{g \in \mathcal{G}} \mathbb{P}(G = g \mid Y_{g-1} = x)\, \frac{1}{5} \sum_{t=1}^{5} \text{FS}_{g,t}(x). \tag{A20}$$

The ratio between the two estimands then recovers

$$\frac{\text{RF}(x)}{\text{FS}(x)} = \sum_{g \in \mathcal{G}} \sum_{t=1}^{5} \omega_{g,t}(x)\, \frac{\mathbb{E}[Y_{g+t}(1) - Y_{g+t}(0) \mid Y_{g-1} = x, G = g]}{\mathbb{E}[B_{g+t}(1) - B_{g+t}(0) \mid Y_{g-1} = x, G = g]}, \tag{A21}$$

with weights

$$\omega_{g,t}(x) = \frac{\mathbb{P}(G = g \mid Y_{g-1} = x)\ \mathbb{E}[B_{g+t}(1) - B_{g+t}(0) \mid Y_{g-1} = x, G = g]}{\sum_{j \in \mathcal{G}} \sum_{k=1}^{5} \mathbb{P}(G = j \mid Y_{j-1} = x)\ \mathbb{E}[B_{j+k}(1) - B_{j+k}(0) \mid Y_{j-1} = x, G = j]}. \tag{A22}$$

The weights are positive and sum to one provided unearned income increases with the prize.

Recovering intensive-margin income effects To isolate the intensive-margin earnings response, we restrict attention to individuals who would work after winning. We impose that, conditional on pre-win earnings $Y_{g-1} = x$, individuals in cohort g who would be working t years after would have experienced the same counterfactual earnings evolution as later winners who are observed to be working in year $g + t$:

$$\begin{aligned}\mathbb{E}[Y_{g+t}(0) - Y_{g-1}(0) \,|\, Y_{g-1} = x,\, G = g,\, Y_{g+t}(1) > 0] \\ = \mathbb{E}\left[Y_{g+t}(0) - Y_{g-1}(0) \,|\, Y_{g-1} = x,\, G > g + t,\, Y_{g+t}(0) > 0\right].\end{aligned} \tag{A23}$$

Two features of our setting make this assumption plausible. First, conditioning on pre-win earnings absorbs most systematic differences in earnings capacity and labor-market attachment. Second, the extensive-margin response is small, so conditioning on being employed at $g + t$ selects nearly the same individuals in the treated and control groups. Together, these features substantially mitigate concerns about selection.

Under (A23), the intensive-margin reduced form becomes

$$\begin{aligned}\text{RF}^{int}_{g,t}(x) &\equiv \mathbb{E}[Y_{g+t} - Y_{g-1} \mid Y_{g-1} = x, G = g, Y_{g+t} > 0] \\ &\quad - \mathbb{E}[Y_{g+t} - Y_{g-1} \mid Y_{g-1} = x, G > g + t, Y_{g+t} > 0] \\ &= \mathbb{E}[Y_{g+t}(1) - Y_{g+t}(0) \mid Y_{g-1} = x, G = g, Y_{g+t}(1) > 0]\,.\end{aligned}$$

The corresponding intensive-margin first stage recovers

$$\text{FS}^{int}_{g,t}(x) = \mathbb{E}[B_{g+t}(1) - B_{g+t}(0) \mid Y_{g-1} = x, G = g, Y_{g+t}(1) > 0]\,. \tag{A24}$$

Our assumption of constant income effects conditional on Y_{g-1} implies that

$$Y_{g+t}(1) = Y_{g+t}(0) + \frac{\eta(x)}{1-\tau}\left(B_{g+t}(1) - B_{g+t}(0)\right),$$

for individuals who work after winning. Substituting this into the intensive-margin reduced form $\text{RF}^{int}_{g,t}(x)$, we obtain

$$\frac{\text{RF}^{int}_{g,t}(x)}{\text{FS}^{int}_{g,t}(x)} = \frac{\eta(x)}{1-\tau}. \tag{A25}$$

Decomposing the total earnings response. To decompose the total earnings effect per additional NOK of unearned income into the intensive- and extensive margin contributions, it is useful to introduce notation for being employed, $E_{g+t} = \mathbb{1}[Y_{g+t} > 0]$ and define the extensive-margin reduced form estimand by

$$\text{RF}^{ext}_{g,t}(x) \equiv \mathbb{E}[E_{g+t} - E_{g-1} \mid Y_{g-1} = x, G = g] - \mathbb{E}[E_{g+t} - E_{g-1} \mid Y_{g-1} = x, G > g + t]. \tag{A26}$$

Under the parallel trends assumption and the additional assumption that employment decreases in lottery winnings, the extensive-margin reduced form estimand recovers:

$$\text{RF}^{ext}_{g,t}(x) = -\,\mathbb{P}(Y_{g+t}(0) > 0,\, Y_{g+t}(1) = 0 \mid Y_{g-1} = x, G = g). \tag{A27}$$

By the law of total expectations, the total earnings effect can be written as:

$$
\begin{aligned}
&\overbrace{\mathbb{E}[Y_{g+t}(1) - Y_{g+t}(0) \mid Y_{g-1} = x, G = g]}^{\text{total effect}=\text{RF}_{g,t}(x)} \\
&= \overbrace{\underbrace{\mathbb{P}(Y_{g+t}(1) > 0 \mid Y_{g-1} = x, G = g)}_{\text{employment share } = \mathbb{P}(Y_{g+t} > 0 | Y_{g-1} = x, G = g)} \times \underbrace{\mathbb{E}[Y_{g+t}(1) - Y_{g+t}(0) \mid Y_{g-1} = x, G = g, Y_{g+t}(1) > 0]}_{\text{intensive margin response}=\text{RF}^{int}_{g,t}(x)}}^{\text{intensive margin component}} \\
&\underbrace{\overbrace{- \mathbb{P}(Y_{g+t}(1) = 0, Y_{g+t}(0) > 0 \mid Y_{g-1} = x, G = g)}^{\text{extensive margin response}=\text{RF}^{ext}_{g,t}(x)} \times \mathbb{E}[Y_{g+t}(0) \mid Y_{g-1} = x, G = g, Y_{g+t}(1) = 0, Y_{g+t}(0) > 0]}_{\text{extensive margin component}} .
\end{aligned}
$$

Since there is only one unknown quantity in the expression, we can solve for it to obtain the average $Y_{g+t}(0)$ for those who would stop working if they won:

$$
\mathbb{E}[Y_{g+t}(0) \mid Y_{g-1} = x, G = g, Y_{g+t}(1) = 0, Y_{g+t}(0) > 0] = \frac{\text{RF}_{g,t}(x) - p_{g,t}(x)\text{RF}^{int}_{g,t}(x)}{\text{RF}^{ext}_{g,t}(x)} \tag{A28}
$$

where the employment share is denoted by $p_{g,t}(x) \equiv \mathbb{P}(Y_{g+t} > 0 \mid Y_{g-1} = x, G = g)$. Dividing the total effect by the first stage then yields:

$$
\frac{\text{RF}_{g,t}(x)}{\text{FS}_{g,t}(x)} = \underbrace{\frac{p_{g,t}(x)}{\text{FS}_{g,t}(x)} \times \text{RF}^{int}_{g,t}(x)}_{\text{intensive-margin component}} + \overbrace{\frac{\text{RF}_{g,t}(x) - p_{g,t}(x)\text{RF}^{int}_{g,t}(x)}{\text{FS}_{g,t}(x)}}^{\text{extensive-margin component}} . \tag{A29}
$$

Since all the quantities in the expression are either identified or functions of the data, the decomposition is identified.

E Derivation of excess burden and revenue-maximizing tax rates

E.1 Excess burden

To derive an expression for the excess burden, it is useful to start by noting that

$$\frac{dDWL}{d\tau} = Y - \sum_{k=1}^{K} \tau_k P_k \frac{\partial C_k^c}{\partial \tau} - Y - (\tau + \tau_w)\frac{\partial Y^c}{\partial \tau},$$
$$= -\left(\sum_{k=1}^{K} \tau_k P_k \frac{\partial C_k^c}{\partial \tau} + (\tau + \tau_w)\frac{\partial Y^c}{\partial \tau} \right),$$

where the first equality uses the envelope theorem. It is well-known from duality theory that the following equality between the compensated and uncompensated demand functions holds.

$$C_k^c(\tau_1, ..., \tau_K, \tau, \bar{V}) = C_k^u(\tau_1, ..., \tau_K, \tau, E(\tau_1, ..., \tau_K, \tau, \bar{V})),$$

where $E(.)$ is the expenditure function defined in equation (29). Two-step budgeting implies that,

$$C_k^c(\tau_1, ..., \tau_K, \tau, \bar{V}) = C_k(\tau_1, ..., \tau_K, I(\tau_1, ..., \tau_K, \tau, E(\tau_1, ..., \tau_K, \tau, \bar{V}))).$$

Taking the derivative of the compensated demand function with respect to τ thus yields,

$$\frac{\partial C_k^c}{\partial \tau} = \frac{\partial C_k^u}{\partial I}\left(\frac{\partial I}{\partial \tau} + \frac{\partial I}{\partial R}\frac{\partial E}{\partial \tau} \right).$$

By recalling that

$$I(\tau_1, ..., \tau_K, \tau, R) = (1-\tau)Y^u(\tau_1, ..., \tau_K, \tau, R) + R,$$

we obtain

$$\frac{\partial I}{\partial \tau} + \frac{\partial I}{\partial R}\frac{\partial E}{\partial \tau} = -Y + (1-\tau)\frac{\partial Y^u}{\partial \tau} + \left[(1-\tau)\frac{\partial Y^u}{\partial R} + 1 \right]\frac{\partial E}{\partial \tau},$$
$$= -Y + (1-\tau)\frac{\partial Y^u}{\partial \tau} + \left[(1-\tau)\frac{\partial Y^u}{\partial R} + 1 \right] Y,$$
$$= -(1-\tau)\frac{\varepsilon^u Y}{1-\tau} + \eta Y,$$
$$= -\varepsilon^c Y,$$

where the second-to-last equality uses the definition of the uncompensated elasticity and the last equality uses the Slutsky equation.

This implies that we can write the sum in the expression for the MDWL as

$$\sum_{k=1}^{K} \tau_k P_k \frac{\partial C_k^c}{\partial \tau} = -\varepsilon^c Y \sum_{k=1}^{K} \tau_k P_k \frac{\partial C_k}{\partial I},$$
$$= -\varepsilon^c Y \frac{\tilde{\tau}}{1+\tilde{\tau}},$$

where the second equality uses the definition of $\tilde{\tau}$ from equation (28) and equation (A30) below.

The marginal deadweight loss can now be written as

$$MDWL = \varepsilon^c Y \frac{\tilde{\tau}}{1+\tilde{\tau}} + (\tau + \tau_w)\frac{\varepsilon^c Y}{1-\tau},$$

which, after rearranging, yields the expression in the numerator of equation (31). To derive the marginal (compensated) tax revenue, we note that

$$\begin{aligned}\frac{\partial TR^c}{\partial \tau} &= \sum_{k=1}^{K} \tau_k P_k \frac{\partial C_k^c}{\partial \tau} + Y + (\tau + \tau_w)\frac{\partial Y^c}{\partial \tau}, \\ &= Y - MDWL.\end{aligned}$$

The expression in the denominator of equation (31) follows immediately.

E.2 Revenue-maximizing tax rate

By standard arguments, increasing the top-income tax rate by $d\tau$ is equivalent to the marginal tax rate changing by $d\tau$ and the virtual transfer changing by $d\tau\bar{Y}$ in a linear tax system. The effect on government revenue can, therefore, be expressed as,

$$\frac{dTR}{d\tau_{\text{TOP}}} = \sum_{k=1}^{K} \tau_k P_k \left(\frac{\partial C_k^u}{\partial \tau} + \frac{\partial C_k^u}{\partial R}\bar{Y}\right) + (Y^u - \bar{Y}) + (\tau + \tau_w)\left(\frac{\partial Y^u}{\partial \tau} + \frac{\partial Y^u}{\partial R}\bar{Y}\right).$$

Using the demand functions in equation (27), we obtain,

$$\frac{\partial C_k^u}{\partial \tau} + \frac{\partial C_k^u}{\partial R}\bar{Y} = \frac{\partial C_k^u}{\partial I}\left(-Y^u + \bar{Y} + (1-\tau)\left(\frac{\partial Y^u}{\partial \tau} + \frac{\partial Y^u}{\partial R}\bar{Y}\right)\right).$$

Substituting this back into the expression above, we get that

$$\begin{aligned}\frac{dTR}{d\tau_{\text{TOP}}} &= \left(-(Y^u - \bar{Y}) + (1-\tau)\left(\frac{\partial Y^u}{\partial \tau} + \frac{\partial Y^u}{\partial R}\bar{Y}\right)\right)\sum_{k=1}^{K} \tau_k P_k \frac{\partial C_k^u}{\partial I} \\ &\quad + (Y^u - \bar{Y}) + (\tau + \tau_w)\left(\frac{\partial Y^u}{\partial \tau} + \frac{\partial Y^u}{\partial R}\bar{Y}\right).\end{aligned}$$

Using the defintion of $\tilde{\tau}$ from equation (28), we have that,

$$1 + \tilde{\tau} = \frac{\sum_{k=1}^{K}(1+\tau_k)P_k \frac{\partial C_k}{\partial I}}{\sum_{k=1}^{K} P_k \frac{\partial C_k^u}{\partial I}} = \frac{1}{\sum_{k=1}^{K} P_k \frac{\partial C_k^u}{\partial I}},$$

where the second equality follows since the budget constraint ensures that $\sum_{k=1}^{K}(1+\tau_k)P_k \partial C_k^u/\partial I = 1$. This means that

$$\sum_{k=1}^{K} P_k \frac{\partial C_k^u}{\partial I} = \frac{1}{1+\tilde{\tau}}.$$

which implies,

$$\sum_{k=1}^{K} \tau_k P_k \frac{\partial C_k^u}{\partial I} = \frac{\tilde{\tau}}{1+\tilde{\tau}}. \tag{A30}$$

Using this, the revenue effect can be written as,

$$\begin{aligned}
\frac{dTR}{d\tau_{\text{TOP}}} &= \left(-(Y^u - \bar{Y}) + (1-\tau)\left(\frac{\partial Y^u}{\partial \tau} + \frac{\partial Y^u}{\partial R}\bar{Y}\right)\right)\frac{\tilde{\tau}}{1+\tilde{\tau}} \\
&\quad + (Y^u - \bar{Y}) + (\tau + \tau_w)\left(\frac{\partial Y^u}{\partial \tau} + \frac{\partial Y^u}{\partial R}\bar{Y}\right), \\
&= -(Y^u - \bar{Y})\frac{\tilde{\tau}}{1+\tilde{\tau}} + (Y^u - \bar{Y}) \\
&\quad + \left((\tau + \tau_w) + (1-\tau)\frac{\tilde{\tau}}{1+\tilde{\tau}}\right)\left(\frac{\partial Y^u}{\partial \tau} + \frac{\partial Y^u}{\partial R}\bar{Y}\right), \\
&= \frac{Y^u - \bar{Y}}{1+\tilde{\tau}} + \left(\tau_w + \frac{\tau + \tilde{\tau}}{1+\tilde{\tau}}\right)\left(\frac{\partial Y^u}{\partial \tau} + \frac{\partial Y^u}{\partial R}\bar{Y}\right)
\end{aligned}$$

From the definitions of the uncompensated elasticity and income effect, we obtain that,

$$\frac{\partial Y^u}{\partial \tau} = -\frac{Y^u \varepsilon^u}{1-\tau}, \qquad \frac{\partial Y^u}{\partial R} = \frac{\eta}{1-\tau}.$$

Substituting this into the expression above,

$$\frac{dTR}{d\tau_{\text{TOP}}} = \frac{Y^u - \bar{Y}}{1+\tilde{\tau}} + \frac{\tau_w(1+\tilde{\tau}) + \tau + \tilde{\tau}}{(1+\tilde{\tau})(1-\tau)}\left(-\varepsilon^u Y^u + \eta \bar{Y}\right)$$

At the revenue-maximizing tax rate, it must be the case that further increases in the top-income tax rate do not affect revenue. Thus,

$$\mathbb{E}\left[\frac{dTR}{d\tau_{\text{TOP}}} \mid Y \geq \bar{Y}\right] = \mathbb{E}\left[\frac{Y^u - \bar{Y}}{1+\tilde{\tau}} + \frac{\tau_w(1+\tilde{\tau}) + \tau_{\text{TOP}} + \tilde{\tau}}{(1+\tilde{\tau})(1-\tau_{\text{TOP}})}\left(-\varepsilon^u Y^u + \eta \bar{Y}\right) \mid Y \geq \bar{Y}\right] = 0,$$

where the expectation is taken across individuals in the top-income tax bracket. Using the definitions from equation (35) and the assumption that $\tilde{\tau}$ is constant across individuals, we obtain,

$$\begin{aligned}
\frac{\mathbb{E}[Y^u - \bar{Y} \mid Y^u \geq \bar{Y}]}{1+\tilde{\tau}} + \frac{\tau_w(1+\tilde{\tau}) + \tau_{\text{TOP}} + \tilde{\tau}}{(1+\tilde{\tau})(1-\tau_{\text{TOP}})}\left(-\bar{\varepsilon}^u\,\mathbb{E}[Y^u \mid Y^u \geq \bar{Y}] + \bar{\eta}\bar{y}\right) &= 0, \\
\frac{(\alpha-1)\bar{y}}{1+\tilde{\tau}} + \frac{\tau_w(1+\tilde{\tau}) + \tau_{\text{TOP}} + \tilde{\tau}}{(1+\tilde{\tau})(1-\tau_{\text{TOP}})}\left(-\bar{\varepsilon}^u \alpha \bar{y} + \bar{\eta}\bar{y}\right) &= 0, \\
(\alpha - 1) + \frac{\tau_w(1+\tilde{\tau}) + \tau_{\text{TOP}} + \tilde{\tau}}{1-\tau_{\text{TOP}}}\left(-\bar{\varepsilon}^u \alpha + \bar{\eta}\right) &= 0
\end{aligned}$$

where the second line uses the defintion of α and the third multiplies both sides by $(1+\tilde{\tau})/\bar{Y}$. Rearranging, we obtain,

$$\frac{\tau_w(1+\tilde{\tau}) + \tau_{\text{TOP}} + \tilde{\tau}}{1-\tau_{\text{TOP}}} = \frac{\alpha - 1}{\bar{\varepsilon}^u \alpha - \bar{\eta}}.$$

Solving for τ_{TOP}, we arrive at

$$\tau_{\text{TOP}} = \frac{\alpha - 1 - (\tilde{\tau} + \tau_w + \tau_w\tilde{\tau})\left(\alpha\bar{\varepsilon}^u - \bar{\eta}\right)}{\alpha - 1 + (\alpha\bar{\varepsilon}^u - \bar{\eta})}.$$

F Dynamic labor supply model

In this appendix, we show how the earnings and unearned income responses to lottery winnings in a dynamic model reflects the income effects defined in the static model in Section 4

We follow Blomquist (1985), Blundell and Walker (1986), and MaCurdy (1983) and consider a standard dynamic life-cycle model. For simplicity, we abstract from non-linear labor income taxation, assume that the return on capital is fixed over time and focus on the case where there is only one consumption good, one margin of labor supply and one asset per period.

Consider an individual who won ξ dollars at time t_0. At time t_0, individuals solve:

$$\max_{\{C_t, Y_t, A_{t+1}\}_{t=t_0}^{T}} \sum_{t=t_0}^{T} \beta^t U(C_t, Y_t), \quad \text{subject to} \quad C_t + A_{t+1} = (1-\tau)Y_t + (1+r)A_t + \mathbb{1}[t = t_0]\xi,$$

where A_t denote assets held at the beginning of period t. The solution to the problem depends on the lottery winnings, so we write $C_t(\xi), Y_t(\xi)$ and $A_{t+1}(\xi)$ for all $t \geq t_0$. The following first-order condition that determines labor supply in period t is:

$$\frac{\partial U}{\partial C}(1-\tau) + \frac{\partial U}{\partial Y} = 0,$$

where the derivatives are evaluated for consumption $C_t = (1-\tau)Y_t(\xi) + B_t(\xi)$ with $B_t(\xi) \equiv (1+r)A_t(\xi) - A_{t+1}(\xi) + \mathbb{1}[t = t_0]\xi$ and earnings $Y = Y_t(\xi)$. This means we can write optimal labor earnings as

$$Y_t(\xi) = Y(\tau, B_t(\xi)),$$

where

$$Y(\tau, B) \equiv \arg\max_{Y} U((1-\tau)Y + B, Y).$$

The time $t \geq t_0$ earnings and unearned-income effect of winning an additional dollar is then given by

$$\frac{dY_t}{d\xi} \quad \text{and} \quad \frac{dB_t}{d\xi}.$$

By the chain rule and the definition of $Y(\tau, B)$,

$$\frac{dY_t}{d\xi} = \frac{\partial Y}{\partial B}\frac{dB_t}{d\xi} = \frac{\eta}{1-\tau}\frac{dB_t}{d\xi},$$

where η is the same parameter as in the static model in Section 4. Thus, the dynamic model implies the income effect is related to the earnings and unearned income effect of lottery winnings through,

$$\frac{dY_t/d\xi}{dB_t/d\xi} = \frac{\eta}{1-\tau}.$$