Statistisk sentralbyrå
Statistics Norway

A practical framework for behavioral microsimulation using external evidence

Zhiyang Jia, Snorre Skagseth, Thor O. Thoresen and Trine E. Vattø

TALL SOM FORTELLER

DISCUSSION PAPERS

1034

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 Statistics Norway staff. The views and conclusions in this document are those of the authors.

Published: February 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

When structural labor supply models are unavailable or impractical, the paper proposes a practical method for incorporating behavioral responses into static taxbenefit microsimulation models. Using elasticity estimates from the literature, the proposed external evidence procedure offers a flexible and transparent tool for incorporating behavioral effects into microsimulation. We distinguish between responses in income both at the intensive and extensive margins. The framework is implemented within the Norwegian LOTTE microsimulation system and illustrated through two policy reforms: (i) a five percentage-point increase in the two top brackets of the labor income tax and (ii) the introduction of a work-related income deduction. Empirical results show that behavioral adjustments substantially offset mechanical revenue effects for high-income tax increases, while the work deduction generates positive participation responses but amplifies revenue losses due to phase-out effects.

Keywords: microsimulation, tax-benefit model, behavioral response, quasi-experimental evidence

JEL Classification: C53, H24, H31

Acknowledgements: This work has been supported by the Norwegian Ministry of Finance. We thank Oddbjørn Raaum, Eric Toder, Lars J. Kirkebøen and participants of Skatteforum 2025 and the 2025 IIPF Congress (Nairobi) for useful comments.

Addresses:
Zhiyang Jia, Statistics Norway, Research Department. E-mail: jia@ssb.no.
Snorre Skagsethm, Statistics Norway, Research Department. E-mail: ssn@ssb.no
Thor O. Thoresen, Statistics Norway, Research Department. E-mail: tot@ssb.no
Trine E. Vattø, Statistics Norway, Research Department. E-mail: tev@ssb.no

Sammendrag

Ulike typer mikrosimuleringsmodeller er utviklet for å gi beslutningsstøtte i utviklingen av skattepolitikken. For norske beslutningstakere er det for dette formålet utviklet en familie av mikrosimuleringsmodeller kjent som LOTTE-modellene. LOTTE-Skatt brukes til å beregne de umiddelbare, direkte effektene av endringer i inntekts- og formuesskatt på skatteproveny og inntektsfordeling, uten å ta hensyn til atferdsresponser. LOTTE-Arbeid bygger en strukturell arbeidstilbudsmodell som anvendes til å anslå virkninger av personskatter på arbeidstilbudet.

I dette arbeidet mangler en imidlertid ofte modeller som kan simulere hvordan aktørene endrer atferd som følge av skatteendringer. Artikkelen presenterer en praktisk tilnærming for å inkludere atferdsresponser i skattepolitiske analyser når strukturelle modeller (slik som LOTTE-Arbeid) ikke er tilgjengelige eller ikke er anvendbare. Metoden bygger på bruk av elastisitetsestimater hentet fra eksisterende litteratur, som anvendes til å justere inntektene i den atferdsfrie mikrosimuleringsmodellen (her LOTTE-Skatt) ved endringer i skattesystemet.

Tilnærmingen består av to hovedtrinn: (1) identifisering av relevante responsestimater fra empiriske studier og (2) integrering av disse i modellen. Det skilles mellom intensive og ekstensive marginer. Den intensive marginen deles videre inn i substitusjons- og inntektseffekter, mens den ekstensive marginen omfatter beslutningen om deltakelse i arbeidsmarkedet. Endringer i arbeidsmarkedsdeltakelse beregnes basert på deltakelsesskatten. Et sentralt element i metoden er imputering av “ kontrafaktisk ” inntekt i alternativ tilstand på ekstensiv margin (arbeid versus ikke-arbeid), som estimeres ved hjelp av paneldata og regresjonsmodeller.

Litteraturgjennomgangen viser betydelig variasjon i elastisitetsestimater, men indikerer generelt moderate atferdsresponser i en skandinavisk kontekst. Metoden illustreres ved to skattereformer: en økning på 5 prosentpoeng i de to øverste trinnene i trinnskatten og innføring av et arbeidsfradrag. I illustrasjonene antas elastisiteten på den ekstensive marginen å være 0,2, mens elastisitetene på den intensive marginen settes til 0,15 for den kompenserte effekten og 0,1 for inntektseffekten. Resultatene viser at atferdsresponser kan ha stor betydning for provenyvirkningene av skatteendringer. For eksempel reduserer økningene i trinnskatten den opprinnelige, mekaniske provenyeffekten med 43 prosent.

Metoden gir et fleksibelt og anvendelig verktøy for å inkludere atferdsvirkninger i praktisk skattepolitisk analyse, og kan enkelt tilpasses nye elastisitetsestimater og utvides til andre atferdsmarginer. Artikkelen gir en detaljert beskrivelse av dette verktøyet.

1. Introduction

In the decision-making process of designing a tax-benefit scheme, governments make use of static microsimulation models to produce information about revenue and distributional effects of alternative policies, describing "day after" effects (Bourguignon and Spadaro, 2006; Immervoll et al., 2007; Figari et al., 2015; Creedy et al., 2020). The term microsimulation reflects that models produce predictions of effects on measures such as disposable income, taxes paid and benefits received for each person and/or household in a representative micro dataset. The effects of a policy change are analyzed by comparing its impact on tax revenue and income distribution against the outcomes of a counterfactual scenario without the policy change. Often, simulation results are limited to output from a standard non-behavioral tax-benefit model.

However, since policy changes are known to influence peoples decisions, policy-makers should ideally be provided with predictions that also reflect the expected behavioral effects of the tax policy change. For instance, it is widely recognized that changes in tax policy influence labor supply, with implications for tax revenue and income distribution. The primary contribution of this study is to show how a policy analyst can meet these demands within a static microsimulation framework, suggesting a reduced-form approach for incorporating response estimates in behavioral microsimulation.

Practical microsimulation work is carried out by agencies such as the Institute for Fiscal Studies (IFS) in the U.K., the Joint Committee of Taxation (JCT) in the U.S., and the European Commission of the European Union. In Norway, Statistics Norway has developed the so-called LOTTE-family of microsimulation models, which includes a behavioral module, to be used to describe the labor supply effects of tax policy changes (Jia et al., 2024). The labor supply module builds on a structural discrete choice model (Dagsvik, 1994; Dagsvik et al., 2014; Dagsvik and Jia, 2016), which is a commonly used methodological framework for modeling of labor supply for use in microsimulation (van Soest, 1995; Creedy and Kalb, 2005).

Structural behavioral models are commonly used to estimate labor supply responses in microsimulation. However, such models are not always available or feasible to apply – for instance, when analyzing reforms that target specific subgroups and therefore impose additional modeling requirements. In these cases, rather than disregarding behavioral effects altogether, analysts may draw on empirical evidence from previous studies to approximate expected responses. In this sense, we refer to our framework as an external evidence approach.

This paper presents a practical procedure for integrating behavioral responses in income into static microsimulation, enabling analysts to account for behavioral adjustments without relying on fully specified structural models. The methodological approach consists of two main steps. First, estimates of behavioral effects relevant to the policy change under consideration are identified from the literature. Second, a procedure is developed to integrate these behavioral estimates into a non-behavioral tax-benefit model. The integration involves adjusting the input of the non-behavioral model to account for expected behavioral responses in income to the tax policy change. The approach distinguishes between behavioral effects on the extensive and intensive margins – for the intensive margin, the responses are further decomposed into components due to substitution and income effects.

To illustrate the suggested procedure, we discuss two changes in the Norwegian personal income tax scheme. Our choice of policy examples is motivated by the goal of demonstrating how the procedure accounts for labor supply effects on both the intensive and extensive margins. First, we examine the extent to which increased taxation for high-income individuals reduces their labor supply. To illustrate this, we increase the two top brackets of the labor income tax, i.e., the fourth and fifth brackets of the bracket tax, by five percentage points. The bracket tax is the progressive component of Norway's dual income tax system and consisted of five brackets in 2022.

Second, we examine the effects of a policy change designed to enhance labor market incentives at the extensive margin: the so-called "work deduction". This policy, proposed by an expert group (Ministry of Finance, 2022), offers a full (100%) deduction on labor income up to a maximum of NOK 55,000 (EUR 5,400; USD 5,700).[1] Similar to the Earned Income Tax Credit in the U.S. and the Working Tax Credit in the U.K., the support is gradually phased out as income increases, potentially creating a disincentive at the intensive margin.

Based on these two reforms, we use appropriate estimates of behavioral response to obtain estimates of tax revenue effects and effects on the income distribution. For the increase in the top marginal tax rate, there is a wide collection of applicable intensive margin response estimates among high-income earners, as taxable income responses for this group have been the main focus of many quasi-experimental studies estimating the elasticity of taxable income (ETI) (Saez et al., 2012; Neisser, 2021). The ETI estimate captures how reported income responds to tax changes through various channels, including labor supply adjustment, efforts, timing of income, and tax planning strategies. Under conventional identifying assumptions, these ETI estimates measure intensive margin effects. Because the overall effect of a tax policy change at the intensive margin reflects both substitution and income effects, we present evidence in terms of a decomposition into a compensated elasticity and an income elasticity.

Similarly, the extensive-margin estimate provides information on the number of taxpayers who enter or exit the labor market in response to tax changes. While we address extensive margin effects for both reforms, the introduction of the "work deduction", presented in the second policy example, is expected to produce more pronounced extensive margin effects. Extensive margin tax responses have received substantial attention, for example in explaining differences in total hours worked across countries (Prescott, 2004; Rogerson and Wallenius, 2009). However, relatively few quasi-experimental studies provide estimates of extensive margin responses; see surveys of the literature in Chetty et al. (2013) and Lundberg and Norell (2020).

Developing a procedure for the extensive margin is inherently more complex than the approach used for the intensive margin. This complexity stems from the fact that individuals can only be observed in one state at a time – either participating in the labor market or not. As a result, a key methodological challenge is assigning income to individuals in their alternative, unobserved state. A key source of information for imputing income comes from econometric analyses that exploit the panel structure of the data; specifically, we use information on individuals who transitioned between participation and non-participation between 2019 and 2022.

We are not the first to evaluate policy changes by combining external evidence on behavioral effects and non-behavioral microsimulation models. Previous studies include Thoresen (2004), Carroll and Hrung (2005), Joint Committee on Taxation (2015, 2023), Thoresen et al. (2012), Ollonqvist et al. (2024), and Finansministeriet and Skatteministeriet (2024). As the methodology has not previously been formalized or thoroughly discussed for policy analysis, this study offers a comprehensive outline of its application. This also includes a comparison of the results of using this "external evidence approach" with simulation results from a structural labor supply model.[2] Furthermore, we demonstrate how the approach can be applied to find the top optimal tax rate, along the lines of Saez (2001).

[1] Average exchange rates for 2022 are used: 1 EUR = NOK 10.1 and 1 USD = NOK 9.6.

[2] The labor supply model belongs to the LOTTE-family of models (Dagsvik et al., 2014; Jia et al., 2024)

The paper is structured as follows. Section 2 explores the broader context of the "external evidence approach" within the framework of the microsimulation tool. In particular, we discuss the key component of our approach: the pool of empirical evidence from which external estimates are derived. Section 3 presents the conceptual framework, highlighting the importance of distinguishing between effects on the extensive and intensive margins. In Section 4, we provide a detailed discussion of each step of our method: the tax policy changes chosen to illustrate it, the choice of response estimates given the tax policy changes, and the practical implementation when there are effects on both the extensive and intensive margins. Furthermore, Section 5 demonstrates the effects of selected policy changes, used as examples to illustrate the proposed procedure, on tax revenue and income distribution. Section 6 concludes the paper.

2. Literature review

2.1 Tax-benefit models

In the following we review the literature that frames the studys methodological approach. It first introduces the role of taxbenefit microsimulation models in policy analysis and then surveys empirical evidence on income responses to tax changes, discussing different elasticity concepts.

Policy-makers often make use of detailed tax-benefit microsimulation models when collecting information about effects of policy change. Tax-benefit models abstracts from behavioral effects but is still a key tool to provide information about different policy alternatives implications in the short-term, often referred to as "day after effects". A selection of static microsimulation models include NBER's TAXSIM (the U.S.), the models of The Joint Committee of Taxation and the Congressional Budget Office (Congress of the U.S.), IFS's TAXBEN (the U.K.), Melbourne University's MITTS (Australia), Treasury's TaxWell (New Zealand), Statistics Sweden's FASIT (Sweden), Treasury's Lovmodellen (Denmark) and EUROMOD of the European Commission.

Similarly, the Norwegian non-behavioral tax-benefit model, LOTTE-Skatt, plays a crucial role in Norway's decision-making system (Jia et al., 2024). LOTTE-Skatt is extensively used by the Ministry of Finance in preparation of the budget and is also made available to the parliamentary opposition for developing their own budgets. Moreover, policymakers may also utilize the labor supply model incorporated within the LOTTE microsimulation system – LOTTE-Arbeid. LOTTE-Arbeid module supplements the non-behavioral LOTTE-Skatt model by estimating labor responses using a job choice framework (Jia et al., 2024; Dagsvik et al., 2014; Dagsvik and Jia, 2016).

While the non-behavioral tax-benefit model (LOTTE-Skatt) constitutes a cornerstone of the Norwegian budget process, LOTTE-Arbeid is considerably more demanding to use and is employed less frequently. This study investigates how the model-system can be further developed by integrating a reduced-form behavioral simulation procedure based on estimates on tax responsiveness from the literature.

2.2 Income responses to tax changes

2.2.1 Different methods and elasticity concepts

A wide range of behavioral responses should be considered when estimating the effects of tax policy changes on revenue and income distribution. For example, models developed for the U.S. Congress account for behavioral effects that can be broadly categorized as shifts in the timing, changes in business sector and entity form, portfolio adjustments, consumption patterns, and tax planning or avoidance strategies (Carroll and Hrung, 2005). These behavioral effects are typically accounted for off-model, meaning that computer sub-routines with relevant behavioral response calculators supplement the core tax-benefit simulations (Joint Committee on Taxation, 2015, 2023). The selection of behavioral margins to include has been extensively debated in the U.S. dynamic scoring literature (Mankiw and Weinzierl, 2006; Auerbach, 2005; Gravelle, 2014).[3] However, given our static (day after) perspective,[4] discussions on how to implement short-run behavioral effects – such as those in Joint Committee on Taxation (2015) – are more relevant. In our applications, we restrict attention to responses in labor income. However, similar external evidence approaches can be applied to other behavioral margins and longer-term effects.

This section provides an overview of the literature, highlighting the primary sources of divergence in the estimates. Notably, when authors refer to a labor supply elasticity, they may be referring to different underlying concepts – each capturing behavioral responses in distinct ways. One distinction can be made between macro-level and micro-level elasticities. In a time series study based on U.S. data from 1946 to 2012, Mertens and Montiel Olea (2018) argue that real activity indicators such as GDP, investment, and employment respond significantly to changes in marginal tax rates, estimating short-run macro elasticities around 1.2.

These findings contrast with micro-level studies based on income tax return data, where (uncompensated) ETI estimates typically range from 0.12 to 0.40 (Saez et al., 2012). Mertens and Montiel Olea (2018) suggest that this discrepancy may be due to their more effective control for endogeneity in marginal tax rates, including factors such as tax policy responses to spending and the business cycle, bracket creep, and anticipation effects. However, macro-level studies based on time series data have been criticized for lacking the granular, quasi-experimental identification strategies often employed in micro-level analyses, which may limit the causal interpretation of their estimates.

Whereas the ETI captures how income responds to changes in the net-of-tax rate, the conventional labor supply elasticity in the structural literature measures how hours worked respond to changes in the wage rate, which may be defined in either gross or net terms. For example, the Hausman model of labor supply (Burtless and Hausman, 1978; Hausman, 1981) explains working hours using the net wage, whereas discrete choice labor supply models (van Soest, 1995; Creedy and Kalb, 2005; Dagsvik et al., 2014) typically refer to changes in the gross wage. Despite their conceptual differences, all elasticity estimates are informative about income responses. This means that, as in the present study, they can be used to predict how tax policy changes influence tax revenue and income distribution.

In microsimulation, structural discrete choice models are widely used to incorporate labor supply responses into tax-benefit models. As already mentioned, Statistics Norway's model portfolio includes a structural labor supply model. Similarly, the European Commissions EUROLAB model applies discrete choice methods to predict changes in working hours and participation across EU countries (Narazani et al., 2021). Although widely applied, structural models have been subject to criticism regarding their robustness and specification (LaLonde, 1986; Imbens, 2010). Their stylized nature and sensitivity to wage estimation methods raise concerns about predictive reliability (Löffler et al., 2018).[5]

[3] As the Congressional Budget Office (CBO) and the Joint Committee on Taxation (JCT) often provide estimates of tax revenues and effects of proposed legislation for a longer time horizon.

[4] In contrast to the static perspective of the present study, the Frisch elasticity is relevant in intertemporal or dynamic settings. Defined by holding the marginal utility of wealth constant, the Frisch elasticity measures the transitory effect of wage changes on hours worked, reflecting how individuals allocate labor effort across life-cycle periods depending on the return to work – see Whalen and Reichling (2017) for a review.

[5] See also discussions on structural versus quasi-experimental approaches in Angrist and Pischke (2010), Heckman (2010), and

The approach adopted in this study can be seen as sidestepping methodological debates over alternative response estimates, instead focusing on applying those estimates that are most relevant to the reform under consideration. Since recent decades of research on income responses to tax changes have been dominated by studies employing quasi-experimental methods to estimate such responses, we next review this literature.

2.2.2 Quasi-experimental evidence

After Lindsey (1987) and Feldstein (1995) demonstrated how ETI estimates can be retrieved from using tax reforms, a wide collection of such estimates have been obtained based on reforms in a number of countries. It is fair to say that response estimates from quasi-experiments have become the main source to information on individuals' tax responsiveness. Part of its appeal stems from the ETI being regarded as a sufficient statistic for evaluating the welfare costs of taxation (Feldstein, 1999; Chetty, 2009), as it captures multiple behavioral margins affecting taxable income (Saez et al., 2012; Neisser, 2021). Estimated elasticities are, however, not immutable parameters – they are correlated with contextual factors (Slemrod and Kopczuk, 2002; Kniesner and Ziliak, 2008).

Given that the ETI estimates are identified by variation in the tax treatment caused by quasi-experimental changes in the tax law (tax reforms), and as there has been a worldwide trend towards lower marginal tax rates (Brys et al., 2011; Piketty et al., 2014), a whole range of ETI estimates based on the high-income response can be found in the literature. Since one of our simulations evaluates the effects of changes in the top marginal tax rate, many estimates from the existing literature are relevant to this example. There are also studies which set out to obtain more general estimates of the tax response. For example, Matikka (2018) uses over time changes in general local tax rates across Finland for identification, whereas others obtain broader elasticity estimates by using several tax reforms in the identification, and reforms not only affecting the high-income earners, confer Giertz (2007) and Kleven and Schultz (2014).

Still, there are several econometric challenges involved in the estimation of the ETI (Saez et al., 2012), which could potentially bias the results. A fundamental econometric problem, given the relationship between marginal tax and income, is that they are jointly determined. After Auten and Carroll (1999) and Gruber and Saez (2002), the main empirical strategy for identification of effects is to establish tax change instruments by calculating tax rate differences based on first period income. But this implies that results are vulnerable to effects of mean reversion, and substantial attention has been devoted to controlling for such effects.[6]

Another influential approach to estimating ETI relies on bunching techniques, which exploit the observed clustering of taxpayers at kink points in the tax schedule where marginal tax rates change, see review in Kleven (2016). This method allows researchers to infer behavioral responses to taxation by comparing the observed distribution of income to a counterfactual smooth distribution in the absence of the kink. Seminal work by Saez (2010) demonstrates how this technique can be used to estimate ETI by analyzing excess mass at kink points in U.S. tax data. Similarly, Chetty et al. (2011) extend the bunching framework by accounting for optimization frictions, showing that the presence of such frictions can attenuate observed bunching and thus bias ETI estimates downward if not properly addressed.

Keane (2010).

[6] The standard differences-in-difference IV technique of Gruber and Saez (2002) has been criticized and alternatives have been suggested, see for example Weber (2014) and Burns and Ziliak (2017). However, there is still no clear consensus of which approach is better.

As an illustration of the ETIs sensitivity to context, it has been argued that Scandinavian countries are likely less exposed to tax evasion and avoidance (Kleven, 2014). The reasoning is that wide coverage of third-party reporting in Scandinavia leads to well-developed information trails, which results in a low level of tax evasion.[7] Given this, it may matter for the implementation of a specific tax change, if the decrease in tax leads to, say, more work effort or less tax evasion. If the tax change makes agents move out of the social security category (more work effort), there is a corresponding social security budgetary effect. In comparison, less tax evasion is reflected by increase in the revenue of the personal income tax due to individuals reporting income more correctly, which is often associated with behavioral changes of the self-employed, as demonstrated by Pissarides and Weber (1989).

In any case, a main point is that the behavioral anatomy of the response estimate may matter for selecting the appropriate estimate. However, it can be difficult to know which effects that are involved. For example, some of the initial ETI studies, as Feldstein (1995), found very large response to the large tax cuts of the U.S. 1986 tax reform (TRA86) and attributed the effects to standard supply-side effects, as labor supply, work effort, etc. Others argue that behavioral effects of the 1986-reform were primarily driven by tax avoidance and income shifting (Slemrod, 1996; Gordon and Slemrod, 2000).[8]

Against this background, one may more generally ask to which extent response estimates are transferable across countries. Kleven (2021) questions the use of such evidence for advice on future reforms, as it can only be used for future reforms that exactly replicate or reverse historical reforms. As already discussed, third-party information reporting, which is common in Norway and other Scandinavian countries, limits the extent of behavioral responses. Ultimately, responses are influenced by government policy, including factors such as the focus on tax administration and enforcement, as well as the choice of tax base (Slemrod and Kopczuk, 2002). An implication of this is that practitioners may prioritize estimates derived from the specific economy under study – in our case, evidence from Norway, as well as assigning greater weight to findings from other Scandinavian countries.

Next, an important question concerns response heterogeneity across income levels. For instance, do high-income individuals respond differently compared to the rest of the population? Several studies of the ETI literature suggest that top income earners are more sensitive to tax changes than other taxpayers (Slemrod, 1996; Saez et al., 2012; Le Maire and Schjerning, 2013; Piketty et al., 2014; Harju and Matikka, 2016; Kreiner et al., 2016; Rubolino and Waldenström, 2019, 2020). However, this likely reflects other response margins – beyond working hours and broad wage income – that disproportionately affect individuals with high incomes. The structural labor supply literature refers, for instance, to opposite findings, with larger wage elasticities for low levels of income; see for example Aaberge et al. (1995) and Dagsvik et al. (2014).

Yet another dimension of the empirical evidence is the size of the reform from which the response estimate is obtained. This is a main point of Chetty (2012). Given that there are adjustments costs, the agents re-optimize only if the reform is large enough and elasticities are therefore attenuated by optimization frictions. With respect to estimation of the ETI, when small tax changes are used in the identification, one may observe no or small behavioral changes because adjustment costs outweigh the benefits of re-optimization. This means that an ETI estimate may reflect the size of the particular reform utilized (as a quasi-experiment) in the estimation.

[7] Correspondingly, Neisser (2021), in her survey of the ETI literature, presents separate results for Scandinavian countries.
[8] Nevertheless, there appears to be a general consensus on smaller response estimates than those reported in Feldstein (1995), typically ranging from 0.12 to 0.40 (Saez et al., 2012).

Finally, since the distinction between behavioral effects at the extensive and intensive margins is central to this study, it is worth noting that relatively few studies have examined tax policy effects at the extensive margin using a quasi-experimental research design (Chetty, 2012). However, we note that there is growing attention in the literature to obtaining estimates of this margin.

In the next section, we outline the conceptual framework for applying elasticities within the "external evidence procedure". In Section 4, we then return to the selection of estimates, based on the tax policy questions used to illustrate the approach.

3. Conceptual framework

3.1 Intensive margin

Our "external evidence procedure" is grounded in standard theory of behavioral responses to taxation. We distinguish between the extensive and intensive margins of labor supply. For the intensive margin, we further decompose behavioral responses into substitution effects and income effects. Since these effects often operate in opposite directions and can differ in magnitude, incorporating information on both enhances the informational content.[9] In particular, identifying the substitution effect (the compensated effect) is crucial for evaluating the excess burden (deadweight loss) associated with taxation.

To identify the substitution effect, we exploit a framework in which the uncompensated effect on taxation of labor income is decomposed into a compensated effect and an income effect. The following draws upon the conceptual framework in Gruber and Saez (2002) and Saez et al. (2012), described in more detail in Finansministeriet and Skatteministeriet (2024).

We express the budget constraint of a taxpayer in terms of the marginal tax rate, MTR, and virtual income, R: $y = (1-MTR)\ell + R$, where y is disposable income and ℓ denotes labor earnings. It follows that virtual income, R, is the sum of non-labor income and the difference between paying the marginal tax on labor income, $MRT \cdot \ell$ and the actual tax paid, t. With utility increasing in disposable income and decreasing in labor effort, utility maximization implies a labor supply function that can be expressed in terms of the marginal net-of-tax rate $(1-MTR)$ and virtual income (R), i.e., $\ell = \ell(1-MTR, R)$.

The uncompensated and compensated elasticities of labor supply with respect to the net-of-tax rate are defined as:

$$\varepsilon^u = \frac{\partial \ell}{\partial(1-MTR)} \frac{1-MTR}{\ell}, \tag{3.1}$$

and

$$\varepsilon^c = \frac{\partial \ell^c}{\partial(1-MTR)} \frac{1-MTR}{\ell}, \tag{3.2}$$

where ℓ^c is the compensated labor income. As the income elasticity is defined in terms of change in virtual income, it is given as:

[9] Note that the extensive margin effect is not decomposed into substitution and income effects.

$$\eta = \frac{\partial \ell}{\partial R}\frac{R}{\ell}. \tag{3.3}$$

Next, we employ the following formulation of the Slutsky equation,

$$\frac{\partial \ell}{\partial(1-MTR)} = \frac{\partial \ell^c}{\partial(1-MTR)} + \ell\frac{\partial \ell}{\partial R}. \tag{3.4}$$

When inserting the elasticity expressions (equations (3.1) – (3.3)) into Equation 3.4, we have $\varepsilon^u \frac{\ell}{(1-MTR)} = \varepsilon^c \frac{\ell}{(1-MTR)} + \ell(\eta\frac{\ell}{R})$, which can be rewritten as

$$\varepsilon^u = \varepsilon^c + \eta\frac{(1-MTR)\ell}{R} = \varepsilon^c + \eta', \tag{3.5}$$

where the semi-elasticity $\eta' = \eta\frac{(1-MTR)\ell}{R} = \frac{\partial \ell}{\partial R}(1-MTR)$ is often the estimated parameter in empirical studies, rather than η.

Furthermore, as already seen, ℓ is a function of the slope of the budget line and virtual income, $\ell = \ell(1-MTR, R)$. We take the total differentiation of labor income, and get $d\ell = \frac{\partial \ell}{\partial(1-MTR)}d(1-MTR) + \frac{\partial \ell}{\partial R}dR$. Using equations (3.1) – (3.3), this can expressed in terms of elasticities, as

$$\frac{d\ell}{\ell} = \varepsilon^u\frac{d(1-MTR)}{1-MTR} + \eta\frac{dR}{R}. \tag{3.6}$$

As we aim to define the income effect in terms of actual tax payments, t, and benefits, b, rather than the virtual income R, we make use of the total differential representation of virtual income:

$$dR = \frac{\partial R}{\partial \ell}d\ell + \frac{\partial R}{\partial MTR}dMTR + \frac{\partial R}{\partial t}dt + \frac{\partial R}{\partial b}db = \ell dMTR - dt + db.$$

Using this expression, along with the Slutsky equation (Equation 3.5), we derive from Equation (3.6):

$$\frac{d\ell}{\ell} = \varepsilon^c\frac{d(1-MTR)}{1-MTR} - \eta\frac{(dt-db)}{R} = \underbrace{\varepsilon^c\frac{d(1-MTR)}{1-MTR}}_{\text{substitution effect}} - \underbrace{\eta'\frac{(dt-db)}{(1-MTR)\ell}}_{\text{income effect}}, \tag{3.7}$$

where the first and second terms on the right-hand side represent the substitution and income effects, respectively.

In the following, we use Equation (3.7) to decompose the behavioral effects of the tax policy under consideration. Our approach involves comparing a new policy scenario, labeled 1, to a reference benchmark, labeled 0. Estimates of the income effect from the literature are used as proxies for the semi-elasticity η'. We then reformulate Equation (3.7) as:

$$\frac{\Delta\ell_i}{\ell_{0i}} = \varepsilon^c\frac{(1-MTR_{1i})-(1-MTR_{0i})}{(1-MTR_{0i})} - \eta'\frac{(t_{1i}-t_{0i})-(b_{1i}-b_{0i})}{(1-MTR_{0i})\ell_{0i}}, \tag{3.8}$$

where the effects are calculated at the individual level, denoted by subscript i. MTR_{0i} and MTR_{1i} are the marginal tax rates before and after the tax change and ℓ_{0i} is taxable labor earnings of individual i prior to the change.[10] Equation (3.8) generates an individual adjustment used to adjust pre-tax income to account for intensive margin behavioral responses. In Section 4, we return to a detailed explanation of the procedure based on the policy changes we use for illustration.

3.2 Extensive margin

The extensive margin of labor supply reflects a discrete choice between participation (P) and non-participation (NP). Following Finansministeriet and Skatteministeriet (2024), we derive expressions when starting from the decision problem of a utility-maximizing individual choosing between these two alternatives.

Let $U^P(y^P, \ell^P)$ denote the utility from participating in the labor market, where $y^P = (1 - MTR)\ell^P + R^P$ is disposable income and ℓ^P denotes labor earnings when working. Similarly, let $U^{NP}(y^{NP}, \ell^{NP})$ denote utility when not participating, where $y^{NP} = (1 - MTR)\ell^{NP} + R^{NP}$ is disposable income, ℓ^{NP} refers to labor earnings (typically 0 or close to 0), and R^{NP} represents non-labor income or transfers in the non-participation alternative. The individual chooses to participate if $U^P(y^P, \ell^P) > U^{NP}(y^{NP}, \ell^{NP})$. The probability of participation thus depends on the net-of-tax wage and the level of virtual income, and changes in tax policy can influence the participation decision by altering the relative gains between the two states.

Work incentives at the extensive margin can be captured by the participation tax rate (PTR) (Saez, 2002; Immervoll et al., 2007; Bartels and Pestel, 2016; Kleven, 2022). It measures the net financial disincentive to work, accounting for taxes and benefits. Formally, it is the share of gross earnings that is effectively taxed away when moving from not working to working. This can be expressed as:

$$PTR = \frac{(t^P - b^P) - (t^{NP} - b^{NP})}{\ell^P - \ell^{NP}} = 1 - \frac{y^P - y^{NP}}{\ell^P - \ell^{NP}}. \tag{3.9}$$

It follows that the higher the PTR, the lower the financial gain from working. When PTR is zero, incentives to take up work are undistorted. From Equation (3.9), we also get an extensive margin expression for disposable income, y:

$$y^P - y^{NP} = (\ell^P - \ell^{NP}) + (b^P - b^{NP}) - (t^P - t^{NP}).$$

This highlights that behind the PTR, there is comparison of disposable income in two hypothetical states: participation (P) and non-participation (NP).

At the aggregate level, let P denote the participation rate (the share of working-age individuals employed), and let PTR denote the participation tax rate. The extensive margin elasticity, λ, is the elasticity of the participation rate P with respect to the net-of-tax rate $(1 - PTR)$,

$$\lambda = \frac{dP/P}{d(1 - PTR)/(1 - PTR)}. \tag{3.10}$$

At the individual level, let P_i be the probability that individual i participates, and let ΔP_i denote the change in that probability induced by the tax reform. We can rewrite Equation (3.10) as

$$\Delta P_i = \lambda P \frac{(1 - PTR_{1i}) - (1 - PTR_{0i})}{(1 - PTR_{0i})}. \tag{3.11}$$

[10] The change in Equation (3.8) is symbolized by Δ to indicate individual-level variation.

In Section 4, we demonstrate how the ΔP_i measure can be applied in practice to identify individuals transitioning between participation and non-participation, which in turn allows us to estimate the revenue and distributional impacts of tax policy changes arising from extensive margin responses.

4. Steps of the procedure

4.1 Data used in the tax-benefit simulation

In the remainder of this study, we illustrate how the "external evidence approach" can inform policy-making in practice. As noted earlier, this approach incorporates behavioral effects by adjusting income based on data from a non-behavioral tax-benefit model. This section provides a detailed description of the main components of the approach, including the data and the non-behavioral simulation model, the selection of response estimates from the literature, and the procedure for obtaining counterfactual incomes for the participation and non-participation states. Next, in Section 5, we present the results of the policy changes that serve to illustrate the approach.

We begin by describing the data source used in our empirical illustration. The non-behavioral tax-benefit model LOTTE-Skatt (Jia et al., 2024) consists of three core components: individual-level annual income data, a comprehensive set of tax rules, and a simulation program that applies these rules to each individual record. In practice, the model draws on many of the same data sources as those used in the Income and Wealth Statistics for Households, which provide complete coverage of the Norwegian population (Statistics Norway, 2019). For tax simulation purposes, the most important data source is the Register of Income Tax Returns.

These income tax return records are linked to a wide range of administrative registers using a unique personal identification number. For example, information on pensions and social transfers is obtained from the Norwegian Labour and Welfare Administration, while data on educational attainment are obtained from the Register of the Population's Level of Education. To construct household units within the model, we use a household register based on the national address register, which includes all residential dwellings in Norway.

The tax-benefit model in this application relies on complete population register data from 2022. The same year, 2022, is also used as the tax law reference year in our simulations of alternative tax schemes.

4.2 Tax policy changes

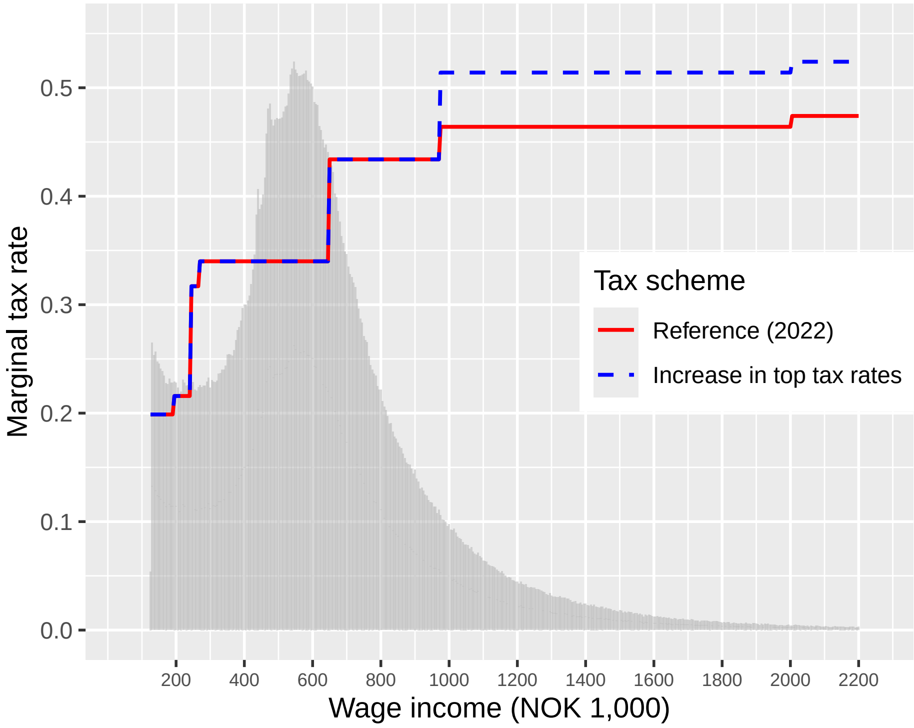
4.2.1 Increased taxes at the top

We demonstrate our procedure by discussing effects of two tax changes. First, we consider a policy change that increases taxation of high-income earners by raising the two top rates of the bracket tax by five percentage-points.[11] In 2022, the fourth bracket, at 46.4%, applied to incomes starting at NOK 969,200 (EUR 96,000; USD 100,000), while the fifth bracket, at 47.4 %,[12] applied to incomes starting at NOK 2 million (EUR 200,000; USD 208,000). With both rates increased by five percentage-points, the new top rates become 51.4% and 52.4%, respectively. Figures 4.1 and 4.2

[11] The term bracket tax was introduced to reflect the expanded number of steps in the revised taxation of labor income in Norway (in place since 2016), compared to the previous two-tier surtax.

[12] This includes the national insurance contribution (8%), the tax on ordinary/net income (22%), and the top rate of the bracket tax scheme (17.4%).

Figure 4.1 **Changes in marginal tax rates from a five-percentage-point rise in the two top bracket tax rates**



Notes: The figure shows marginal tax rates under the reference tax scheme (2022) and after introducing a five-percentage-point increase in the two top bracket tax rates. The horizontal axis shows observed wage income. The gray density curve illustrates the distribution of all wage earners in 2022.

illustrate the effects of these changes on marginal tax rates and participation tax rates, using the 2022 tax law as the benchmark.[13]

4.2.2 Introduction of a work deduction

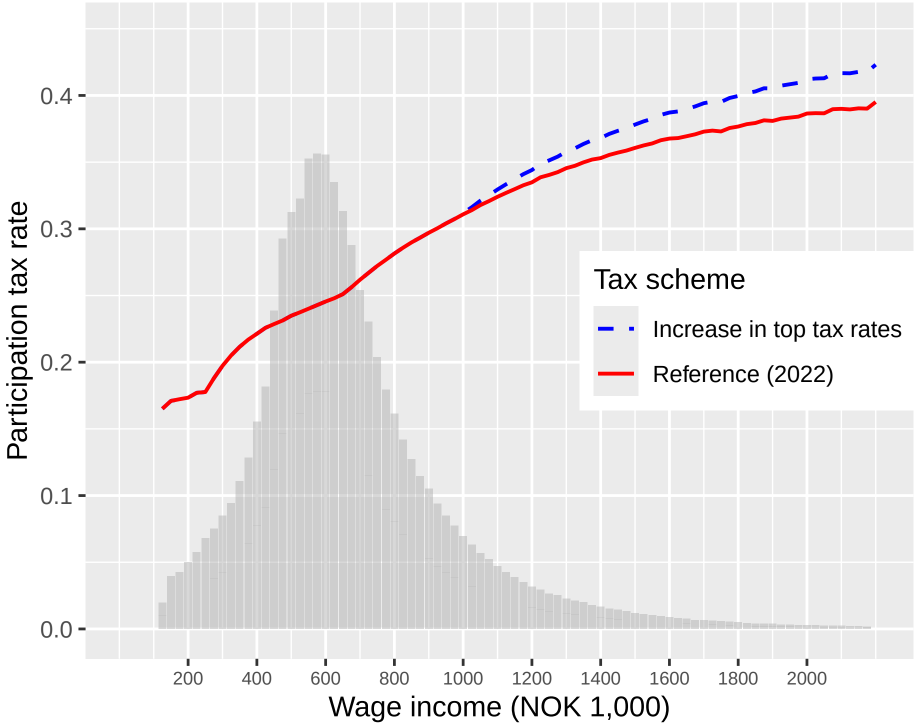
To further illustrate the methodology for deriving extensive margin effects, we examine the impact of the work deduction – a policy designed to promote labor market participation – proposed in a government green paper (Ministry of Finance, 2022). The policy change permits a full (100%) deduction on labor income up to a ceiling of NOK 55,000 (EUR 5,500; USD 5,700). Beyond this threshold, the deduction is gradually phased out. Specifically, the phase-out begins at an income level of NOK 300,000 (EUR 29,700; USD 31,200), with the deduction reduced by 5% for each krone earned above that point. As a result, the deduction is fully eliminated at an income of NOK 1.4 million (EUR 140,000; USD 145,000).[14]

Figures 4.3 and 4.4 illustrate how this reform affects the marginal tax rate and the participation tax rate across the earnings distribution. While the new scheme eliminates income tax entirely for incomes below NOK 55,000 (EUR 5,500; USD 5,700), the effect at the very bottom of the income scale is limited, since the initial tax burden under the reference system (the 2022 tax schedule) was already low.

[13] Given the distinction in the ETI literature between estimates based on broad income (before deductions) and taxable income (after deductions), see e.g., Gruber and Saez (2002) and Neisser (2021), it is worth noting that the change in the bracket tax applies to broad labor income (wage income and the earnings share of self-employment income).

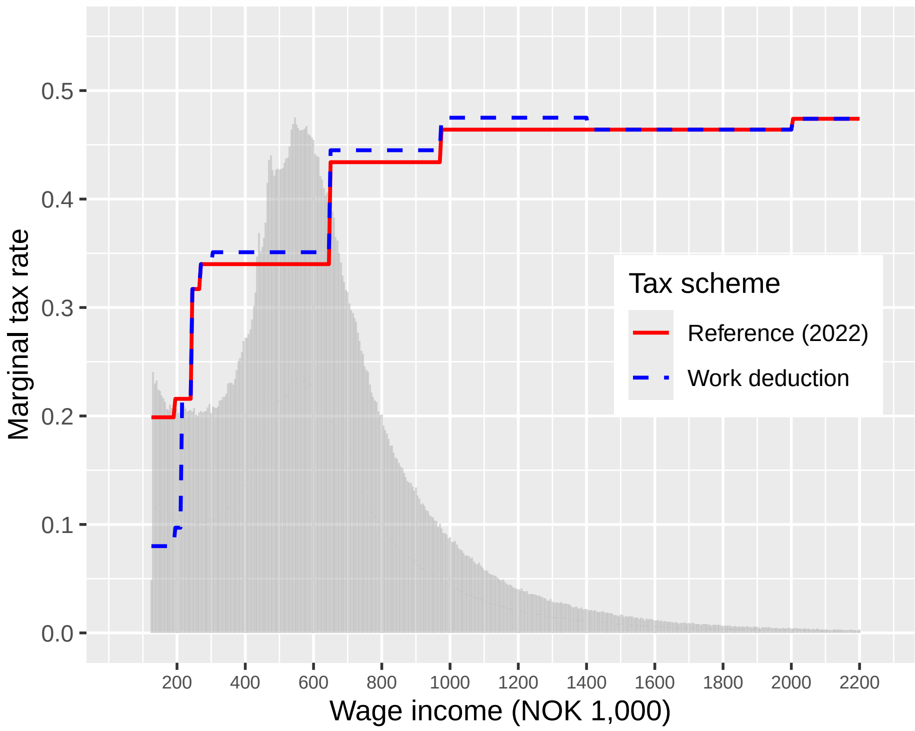
[14] This follows from the calculation: $55,000/0.05+300,000$.

Figure 4.2 **Participation tax rate changes from a five percentage-point increase in the top tax rate**



Notes: The figure shows participation tax rates (PTR) under the 2022 tax scheme and under an alternative scenario with a five percentage-point increase in the two top bracket tax rates; blue dashed line for the latter. For each wage level, the PTR represents the median rate for observed workers who do not receive disability benefit in the non-participation state (we will return to further details about income in non-participation in Section 4.4). The horizontal axis displays observed wage income. The gray curve represents the corresponding density of wage earners in 2022.

Figure 4.3 **Change in marginal tax rates resulting from the work deduction**



Notes: The figure shows marginal tax rates for the benchmark scheme (2022) and an alternative with a work deduction, phased out over the interval from NOK 300,000 to NOK 1,400,000. The gray density curve represents the full population of wage earners in 2022.

Figure 4.4 **Change in the participation tax rate resulting from the work deduction**

Notes: The figure shows participation tax rates (PTR) under the 2022 tax scheme, compared to an alternative with the work deduction. For each wage level, the PTR represents the median rate among observed workers whose alternative to working is not receiving disability benefit. The blue dashed line shows the PTR for an alternative with the work deduction. The horizontal axis displays observed wage income. The gray curve represents the corresponding density of wage earners in 2022.

4.3 Choice of response estimates

4.3.1 Intensive margin estimates

As emphasized, a key component of the "external evidence procedure" is the selection of estimates for both the extensive and intensive margins of behavioral responses. In this section, we present our chosen estimates for the intensive margin, while Section 4.3.2 addresses the selection of response estimates for the extensive margin. At the intensive margin, we particularly draw on the well-established ETI literature (see Section 2). However, as we will show, the literature remains incomplete, and the external validity of existing estimates is limited when applied to settings beyond those in which they were obtained.

As detailed in Section 3, we utilize the Slutsky equation to distinguish between uncompensated, compensated, and income effects. This distinction highlights a key problem: many studies do not clearly separate uncompensated and compensated effects and thereby also provide only limited insight into income effects. The issue arises because the ETI literature lacks a standardized econometric method for differentiating these effects – for example, by identifying income effects – and, as a result, many studies do not report them.

Gruber and Saez (2002) argue that the ETI estimate in their study primarily reflects a compensated effect, assuming the income effect is negligible. In their empirical investigation, Gruber and Saez (2002) differentiate between income and substitution effects by considering several tax reforms and the distance to kinks in the tax scheme for identification. Blomquist and Selin (2010) adopt an approach closer to the structural labor supply literature to identify income effects, which is the same method as used by Thoresen and Vattø (2015). They report that it is challenging to obtain robust estimates of both the income effect and the substitution effect using the same variation in tax changes.

Table 4.1 presents elasticity estimates from a selection of studies we consider most relevant to our setting. In addition to presenting average estimates for the Scandinavian countries (Denmark, Norway, and Sweden), we include results from the ETI literature survey by Neisser (2021).[15] We also refer to estimates of single studies from Norway, Sweden, and Denmark. Some of the included studies provide evidence consistent with the Slutsky decomposition.

The literature review in Table 4.1 clearly shows substantial variation in elasticity estimates. However, none of the studies reviewed report very large behavioral responses. With the exception of Hansson (2007), most studies point to uncompensated elasticities around 0.2 and lower. Relatively small response estimates are further supported by a trend toward smaller responses in working hours observed in several studies; see, for example, Blau and Kahn (2007) and Heim (2007) for U.S. evidence, and Jia et al. (2025) for Norwegian results.[16] Jia et al. (2025) use repeated cross-sectional data and estimate a structural discrete choice model for each year from 1997 to 2019. For married females, the wage elasticity of working hours at the intensive margin decreases from approximately 0.3 in 1997 to about 0.2 in 2019. In contrast, the responsiveness of married males at the intensive margin has remained relatively stable, at around 0.1 throughout the period.

As previously discussed, since both tax policy changes under consideration affect the upper end of the income distribution, it is crucial to obtain estimates that align with this segment.[17] The ETI estimates from Norwegian studies (Aarbu and Thoresen, 2001; Thoresen and Vattø, 2015; Vattø, 2020) are derived by exploiting variation in tax treatment from specific Norwegian tax reforms – the 1992-reform and the 2006-reform – which primarily involved reductions in top marginal tax rates. While these studies examine past periods, we contend that their findings are relevant for the current analysis: Aarbu and Thoresen (2001) find a main (overall) ETI estimate of 0.2 for the uncompensated effect, whereas both Thoresen and Vattø (2015) and Vattø (2020) report somewhat smaller estimates, below 0.1 and around 0.1, respectively. Table 4.1 also reports the income elasticity for married females from Thoresen and Vattø (2015), which is small: -0.02.[18]

In summary, considering the tax policy changes and the available empirical evidence, we anticipate relatively low responsiveness among Norwegian taxpayers at the intensive margin. We use estimates of 0.1 for the uncompensated response in income, 0.15 for the compensated effect, and -0.05 for the income effect. These estimates are close to those underlying the Danish simulation model, which is also based on evidence retrieved from the literature, see Finansministeriet and Skatteministeriet (2024).

4.3.2 Extensive margin estimates

Evidence on extensive margin responses is more limited than that for the intensive margin. But recent studies have begun to address this gap, including Alpert and Powell (2020), Gelber et al. (2020), and Bastani et al. (2021). Typically, responsiveness at the extensive margin is assumed

[15] Neisser (2021) differentiates between estimates based on broad income (before deductions) and taxable income. This distinction is not crucial in the present context, as the taxation of labor income and self-employment income under Norway's personal income tax scheme essentially does not involve any deductions, making taxable income nearly identical to broad (or gross) income.

[16] However, Klemm et al. (2018) argue that there is no downward trend in ETI estimates.

[17] As noted in Section 2.2.2, high-income earners tend to respond differently to tax changes compared to middle- and low-income taxpayers, largely due to their greater access to tax planning strategies and avoidance opportunities. Consequently, quasi-experimental studies often find that top income earners are more responsive to tax changes than other segments of the population.

[18] As already noted, quasi-experimental estimates of income effects are subject to substantial uncertainty due to methodological challenges.

Table 4.1 A collection of estimates from quasi-experimental studies analyzing the intensive margin in the literature

Study	Population group/country	Uncomp. effect	Compens. effect	Income effect
Neisser (2021)[a]	Scand. average	0.18		
Kleven and Schultz (2014)	Wage earners/Denmark	0.05		-0.01
Jakobsen and Søgaard (2022)[b]	All/Denmark	0.2/0.01		
Hansson (2007)	Wage earners/Sweden	0.4–0.5		
Gelber (2014)	Male/Female married wage earners/Sweden	0.21/0.23	0.41/0.47	-0.07/-0.05
Miao et al. (2025)	High-income/Sweden	0.16		
Aarbu and Thoresen (2001)	All/Norway	0.2		
Thoresen and Vattø (2015)	Wage earners/Norway	0.06		
	Married female wage earners/Norway	0.04	0.06	-0.02
Vattø (2020)[c]	Wage earners/Norway	0.11		
Berg and Thoresen (2020)	Self-employed/Norway	0.17		
Graber et al. (2022)[d]	Wage earners/Norway	0.23		

Notes: [a]The estimate for broad (before deductions) income is reported. The after-deduction estimate is somewhat smaller, around 0.1. [b]The two estimates reported refer to taxable income (0.2) and broad income (0.01). [c]The conventional ETI estimate is reported, whereas the longer-term estimate from a dynamic panel data model approach is larger (around 0.15). [d]The estimates reported refer to the point estimates of the medium income earner.

to be somewhat larger than at the intensive margin (Blundell et al., 2013). However, some studies suggest smaller estimates – for instance, Kleven (2022) challenges earlier findings that the U.S. Earned Income Tax Credit (EITC) has had significant extensive margin effects. Contrary to the prevailing view, Kleven (2022) concludes that EITC expansions have shown no clear or significant impact on employment.

Chetty et al. (2013) provide a comprehensive review of micro-econometric estimates of labor market participation responses, reporting wage elasticities between 0.13 and 0.43, with an average of 0.25. This corresponds to the mean of the micro evidence, based on a collection of nine studies. Furthermore, the survey by Lundberg and Norell (2020) reports relatively large estimates, reviewing studies that use quasi-experimental methods to assess the elasticity of labor force participation with respect to the financial gain from work. Across 35 studies, they report an average elasticity of 0.36, with estimates generally higher in the United States than in Europe, and somewhat higher for women and older individuals. They suggest that the policy-relevant elasticity for the full population likely lies in the range of 0.1 to 0.2.

Regarding individual studies, it is worth noting that Bastani et al. (2021) find an average participation elasticity of 0.13 for married females in Sweden. They use a significant reform in Sweden's tax and transfer system to estimate the effects on secondary earners. Evidence in Jia et al. (2025) also supports relatively modest responses at the extensive margin, based on repeated estimations of a structural discrete choice model from 1997 to 2019. For married females, the extensive margin wage elasticity declined markedly, from approximately 0.45 in 1997 to around 0.1 in 2019. In contrast, the male extensive margin elasticity remained consistently low throughout the period, at roughly 0.01–0.04.

To allow for a larger effect at the extensive margin than at the intensive margin, we set the extensive margin response to 0.2.[19] This choice aligns with the survey results reported in Chetty et al. (2013) and Lundberg and Norell (2020).

[19]Recall that the extensive margin response is primarily interpreted as a substitution effect in the literature.

Table 4.2 summarizes our chosen elasticity estimates, where the intensive margin uncompensated effect is decomposed into a compensated effect and an income effect. While these are the main estimates used, Appendix B provides robustness checks with respect to alternative response assumptions.

Table 4.2 Choice of response estimates for the empirical illustrations

	Extensive margin (λ)	Intensive margin (ε^u) Compensated effect (ε^c)	Income effect (η')
Elasticity estimate	0.2	0.15	-0.05

Notes: The response estimates are selected based on interpretations of the evidence reported in the literature, conditional on the tax policy changes used in the illustration. We assume that elasticities are homogeneous across the population considered. See Section 3 for further details about the decomposition of the intensive margin response, including the definition of η'.

4.4 Empirical implementation

4.4.1 Income adjustment at the intensive margin

The following section outlines how elasticity estimates are integrated into the taxbenefit microsimulation model to predict behavioral responses of a tax change. We begin by examining the intensive margin behavioral effects, as derived from Equation (3.8). For each participating individual i, we compute the marginal tax rate (MTR) in the baseline and under each policy change by increasing her labor earnings, ℓ_i^P, by a small amount, $\Delta\ell_i^P$ (NOK 100), while holding all other income sources, including her spouse's income, fixed:

$$MTR_i = \frac{t(\ell_i^P + \Delta\ell_i^P) - t(\ell_i^P)}{\Delta\ell_i^P}.$$

Denote pre-reform and post-reform values by MTR_{0i} and MTR_{1i}, respectively. Next, by substituting the marginal tax rates into Equation (3.8), we obtain individual adjustment factors, $\frac{\Delta\ell_i^P}{\ell_i^P}$, using the chosen compensated elasticity, ε^c, and the income effect estimate, η', together with the individual pre-reform and post-reform tax rates, t_{0i} and t_{1i}.

Consider the policy change involving a five percentage-point increase in the marginal tax rates of the fourth and fifth brackets of the bracket tax as an example. It is assumed to affect individuals with income over approximately NOK 970,000 (EUR 96,000; USD 100,000), where the fourth bracket begins.[20] It follows that an adjustment factor is constructed for each individual according to Equation (3.8) and used to adjust pre-tax income for the affected group of individuals. The post-reform tax system is then applied to the adjusted income levels.

4.4.2 Counterfactual income and the extensive margin

As previously noted, estimating effects at the extensive margin presents several empirical challenges, rendering it more complex than deriving intensive margin results. A primary complication arises from the need to determine each individuals income, tax payment, and disposable income under both scenarios – when working and when not working. These values are essential for calculating the participation tax rate for each individual, as given by Equation (3.9). Since individuals can only be observed in one state at a time – either participating in the labor market or not – it is

[20] We neglect that individuals located at kinks may re-optimize due to the tax change (Christiansen et al., 2022). This implicitly rules out possibilities for larger discrete adjustments, as for example implied by the structural discrete choice labor supply model. In a discrete choice model, individuals are not assumed to optimize at the margin; instead, they are assumed to select the discrete alternative that yields the highest utility. We return to simulation results for a discrete choice model in Section 5.3.

necessary to impute their income in the unobserved state, that is, to assign their counterfactual income. Additional details of the procedure are provided in Appendix A, while Appendix B discusses the implications of certain empirical choices based on various sensitivity checks.

To impute counterfactual income for participants and non-participants, we rely on the Income and Wealth Statistics for Households (Statistics Norway, 2019), the main data source for the taxbenefit model, exploiting its panel dimension to generate empirical measures for imputation. Specifically, we focus on individuals who transitioned between labor market participation and non-participation during the period from 2019 to 2022. For instance, counterfactual wage income is estimated using only those individuals who were employed in 2022 but not in 2019, rather than using the entire population of workers observed in 2022. This choice reflects the fact that individuals not working in 2019 are likely a selected group with, on average, lower earnings potential than the general working population. Crucially, this group closely resembles the target population – those who may enter the labor market in response to tax policy changes.

We begin by restricting our sample to individuals of working age, defined as those between 25 and 61 years old. Additionally, we exclude students, retirees, and the self-employed, yielding a final sample of approximately 2.6 million individuals. To classify individuals as either working or not working, we apply wage income thresholds set at 1 G in 2022 (NOK 111,477; EUR 11,000; USD 11,600).[21] In Appendix B, we examine how sensitive the results are to applying alternative thresholds (0.1 G, 0.4 G, or 2 G) for classifying individuals as either working or non-working.

Next, we outline the procedure for estimating market income for non-participants and then describe how income is assigned in the non-participation state for individuals who participate. These counterfactual income levels, together with the observed state, are used to compute PTR's for each individual both before (PTR_{0i}) and after (PTR_{1i}) the tax reform, as seen in Equation (3.11). Furthermore, we outline the approach for identifying individuals who, according to the method, transition between participation and non-participation as a result of the reform. Since we are analyzing the distributional effects of a reform, it is advantageous to allow some individuals to actually change states in response to the tax change, rather than relying solely on individual probabilities.

Counterfactual income for the observed non-working

To impute wages for non-participants, we use a subset of individuals who were non-participants three years prior (2019) but are participating in 2022, and estimate a Mincer-type wage regression with log earnings as the dependent variable:

$$\log \ell_i = \alpha + Z_i \lambda + \mu_i, \tag{4.1}$$

where ℓ_i represents observed wage income in 2022, α is a constant, Z_i is a vector of explanatory variables, and μ_i is the error term. The explanatory variables include potential work experience, education level and field, gender, centrality index, immigration background, and a categorical variable for disability support received in 2019, including the disability benefit and the work assessment allowance[22], discretized into seven dummy variables based on the amount received. Further details on model specifications and estimation results are provided in Table A2 in Appendix A.

This model is used to impute counterfactual wage income (ℓ_i^P) for all observed non-working individuals in 2022. Individual-specific random noise ($\hat{\mu}_i$) is added to the mean predictions to reflect

[21] G refers to the basic amount in the Norwegian national insurance scheme.

[22] Work assessment allowance is a temporary benefit in Norway for people who cannot work because of illness or injury. Its main purpose is to support individuals financially while they get treatment, rehabilitation, or help to return to work, but in practice it is often a first step towards a disability benefit.

individual variation:

$$\widehat{\ell_i^P} = \exp(\widehat{\alpha} + Z_i\widehat{\lambda} + \widehat{\mu_i}), \tag{4.2}$$

where $\widehat{\ell_i^P}$ symbolizes that we obtain imputed measures of labor income.[23]

Furthermore, we determine the level of government support in terms of benefits and transfers, $\widehat{b_i^P}$, for the observed non-working group in the counterfactual state of labor market participation. We assume that all transfers in the participation state equals the observed level of benefits in non-participation (b_i^{NP}), except for the most common means-tested benefits: disability benefit and "social assistance". Disability benefits refers to both disability benefits and the work assessment allowance. Social assistance refers to the combined amount of direct financial assistance from the Norwegian Labour and Welfare Administration (økonomisk sosialhjelp) and housing allowance (bostøtte) from The Norwegian State Housing Bank (Husbanken).

First, we phase out the disability benefit (if observed) based on predicted wage income in participation, $\widehat{\ell_i^P}$, following the rules of the support scheme. As disability benefits are phased out with 66% of additional wage income above 0.4 G (annually), we assume the following relation: $\widehat{b_i^P} = b_i^{NP} - 0.66\max(\widehat{\ell_i^P} - \ell_i^{NP} - 0.4\text{G}, 0)$. Second, we scale down social assistance (if observed). The scaling factor is derived from individuals who were non-participants in 2019 and participants in 2022, conditional on receiving social assistance in 2019. This subset is grouped by a centrality index (as support varies geographically) to compute conditional means of social assistance for 2019 and 2022. The factor, calculated as the ratio of these means, is applied to scale down the observed social assistance transfers for non-participant in the counterfactual state of participation. Further details are provided in Appendix A; see Table A5.

In summary, to calculate the disposable income of the observed non-participants in the counterfactual scenario of participation, we compute $\widehat{y_i^P} = \widehat{\ell_i^P} + k_i + \widehat{b_i^P}(\widehat{\ell_i^P}) - \widehat{t}_i(\widehat{\ell_i^P}, k_i, \widehat{b_i^P})$ based on the predicted levels of labor earnings and benefits, given by $\widehat{\ell_i^P}$ and $\widehat{b_i^P}$, and the corresponding predicted (simulated) tax, $\widehat{t}_i$. k_i is capital income.

Counterfactual income for the observed working population

As already described, non-participation encompasses a variety of support schemes. A substantial share of individuals classified as non-participants in 2022 receive disability benefits (57%) or social assistance (11%). We therefore treat these programs as the primary benefits for individuals who hypothetically leave the labor market and become non-participants.[24] For this group, we employ a two-step procedure to determine their counterfactual benefits, b_i^{NP}. First, to estimate both the likelihood of receiving disability benefits and the associated benefit level, and second, to estimate the likelihood of receiving social assistance and the corresponding amount. The remaining individuals receive neither benefit in the non-participation state, consistent with observed empirical patterns. Finally, we define disposable income in the counterfactual non-participation state.

First, to determine disability benefit eligibility, we estimate a logistic regression model using a subsample of individuals who participated in the labor market in 2019 and were non-participants three years later – in 2022. The dependent variable $D_i \in \{0, 1\}$ indicates whether an individual receives disability benefit. The explanatory variables include all the same variables as used in the wage regression plus the inverse hyperbolic sine of each individual's 2019 wage income. This model

[23] To ensure the model predicts higher wage income for all individuals in the state of participation than in the state of non-participation, we apply left truncation to the predicted wage distribution at 1.1 G.

[24] We do not consider unemployment benefits as an alternative, as this type of income support is typically short-term and less relevant for more permanent participation decisions.

yields the probability of receiving disability benefits in non-participation, $Pr_i(D=1)$. Estimation results are shown in Table A2 in Appendix A. In the following, we denote individual is counterfactual disability benefit recipiency indicator as follows: $\widehat{q}_{di} \sim \text{Bernoulli}\big(\text{Pr}_i(D=1)\big)$. The disability benefit received by individual i,, d_i, is calculated according to the rules of the scheme as 66% of their observed wage income, subject to a maximum equal to 66% of 6 G (NOK 401,317; EUR 39,700; USD 41,800).

Second, individuals predicted not to receive disability benefits in non-participation may instead receive social assistance, s_i. To estimate these amounts, we employ a subsample of the panel data by conditioning on non-receipt of disability benefits and use a two-step correction function approach to assign social assistance to each individual.

In the first step, we model the probability of receiving social assistance, $S_i \in \{0,1\}$, using a logistic regression. In the second step, we estimate the amount received, $\widehat{s}_i$, conditional on receiving, using OLS. We impute the counterfactual social assistance indicator as: $\widehat{q}_{si} \sim \text{Bernoulli}\big(\text{Pr}_i(S=1)\big)$, yielding a final social assistance amount of $\widehat{q}_{si}\widehat{s}_i$. See Section A.3 in the Appendix for more details.

Thus, the counterfactual transfers in non-participation for the observed working population consists of disability benefits and social assistance:[25]

$$\widehat{b_i^{NP}} = (1-\widehat{q}_{di})\,\widehat{q}_{si}\widehat{s}_i + \widehat{q}_{di}\,d_i. \tag{4.3}$$

This formula ensures that individuals assigned to receive disability benefits ($\widehat{q}_{di}=1$) receive no social assistance. Moreover, individuals not receiving disability benefit ($\widehat{q}_{di}=0$) receive social assistance only if $\widehat{q}_{si}=1$. If both $\widehat{q}_{si}=\widehat{q}_{di}=0$, the individual receives no support in the counterfactual state.[26]

Correspondingly, the counterfactual disposable income in non-participation for the observed working population is then given by,

$$\widehat{y_i^{NP}} = \ell_i^{NP} + k_i + \widehat{b_i^{NP}} - t_i(\ell_i^{NP}, k_i, \widehat{b_i^{NP}}). \tag{4.4}$$

where labor income, ℓ_i^{NP} is set to zero, k_i is observed capital income, $\widehat{b_i^{NP}}$ represents transfers defined as in Equation (4.3) and $t_i(\ell_i^{NP}, k_i, \widehat{b_i^{NP}})$ are computed taxes assigned to the counterfactual income levels.

Finally, for each individual, whether working or not, this procedure imputes counterfactual disposable income for the unobserved labor market state. For workers, we observe y^P and impute y^{NP}, whereas for non-participants, we impute y^P and observe y^{NP}. This allows us to compute individual participation tax rates, as defined in Equation (3.9), under both the reference tax schedule rules (PTR_{0i}) and the new tax rules (PTR_{1i}) for all individuals.

Selection of switching individuals and revenue effects

As noted above, it is practically advantageous to specifically select individuals to move between participation and non-participation in response to the tax changes. Assigning switches at the individual level – that is, identifying which individuals in the dataset move from participation to non-participation or vice versa in response to the policy change – allows for a detailed characterization of distributional outcomes.

To select individuals who switch, we utilize the variable ΔP_i, defined as the tax reform induced

[25] In addition, there are other general transfers, such as child benefit.

[26] The lack of income support for some individuals in the non-participation state is in line with empirical patterns in the data.

change in the participation probability in Equation (3.11). A given reform can induce entries or exits depending on the sign of ΔP_i, which in turn reflects the change in the participation tax rate, see Equation (3.11). If $\Delta P_i > 0$, this reflects an increase in the probability of participating induced by a decline in the participation tax rate: individuals who were previously not working (NP) may enter employment (P), while current workers do not respond. Conversely, if $\Delta P_i < 0$, the probability of participation decreases due to a rise in the participation tax rate: only those who were working before the reform may switch to non-participation (NP), and non-workers remain unaffected. Table 4.3 summarizes how these potential status changes depend on pre-reform labor market status and the sign of ΔP_i.[27]

Table 4.3 Potential labor market status change for a tax reform

	Working (P)	Not working (NP)
$\Delta P_i \geq 0$	No response	May switch to P with probability ΔP_i
$\Delta P_i < 0$	May switch to NP with probability $\lvert\Delta P_i\rvert$	No response

Furthermore, for each non-working individual i in our sample, we run a Bernoulli trial with parameter $\max(0, \Delta P_i)$ to determine whether the individual will enter into participation, i.e., $f_i = 1$ or not $f_i = 0$. Then, the sum $\sum_{NP} f_i$ provides an estimate of the number of individuals who start working as a result of the reform. Similarly, for each working individual, i, under the prereform tax scheme, we run a Bernoulli trial with parameter $|\min(0, \Delta P_i)|$ to determine whether the individual will stop working, $g_i = 1$, or not, $g_i = 0$.

The net increase in labor participation is then simply the difference between those who enter and those who exit: $\sum_{NP} f_i - \sum_P g_i$. Correspondingly, the net revenue change (change in taxes minus benefits), ΔM, is given by:

$$\Delta M = \sum_{d_i=1} \left[t_1(\ell_i^P, k_i, b_i^P) - t_0(\ell_i^{NP}, k_i, b_i^{NP}) \right] - \left[b_1(\ell_i^P, k_i, b_i^P) - b_0(\ell_i^{NP}, k_i, b_i^{NP}) \right]. \tag{4.5}$$

Obviously, the set of individuals who switch labor supply status and the associated tax revenue change represent only one realization of several possible outcomes. To account for this uncertainty, we generate a bootstrap sample of size 500, from which we compute the sample mean, median, and confidence intervals for ΔM. The set of individuals who transition is determined by the scenario whose predicted tax revenue change most closely aligns with the sample median.

5. Illustrations of the procedure

5.1 Effects of increased top tax rates

In this section, we demonstrate how the "external evidence approach" can be used to generate empirical estimates of the revenue and distributional effects of policy changes. We also compare our results with simulation outcomes generated by a structural behavioral microsimulation model – the labor supply module for wage earners of the Norwegian microsimulation system LOTTE (Dagsvik et al., 2014; Dagsvik and Jia, 2016; Jia et al., 2024).

First, we refer to the results from applying the "external evidence approach" to analyze the five percentage-point increase in the two top brackets of the bracket tax. The revenue simulation results are presented in Table 5.1. We find that the direct non-behavioral revenue effect of the increased tax rates is approximately NOK 7.9 billion (EUR 0.78 billion; USD 0.82 billion).

[27] Empirically, we restrict PTR_0 and PTR_1 to the interval $[0, 0.95]$ to avoid division by zero in Equation (3.11).

However, this revenue increase is curbed by behavioral adjustments. Table 5.1 shows that, after accounting for behavioral responses, the tax reform generates a net revenue gain of about NOK 4.25 billion (EUR 0.42 billion; USD 0.44 billion). Our estimates indicate that behavioral responses lower total revenue by about NOK 3.6 billion (EUR 0.36 billion; USD 0.38 billion), driven by intensive margin effects accounting for roughly NOK 2.9 billion (EUR 0.29 billion; USD 0.30 billion). The extensive margin effects account for only NOK 0.5 billion (EUR 0.05 billion; USD 0.05 billion), reflecting that approximately 800 individuals transition from participation to non-participation.[28]

Furthermore, the aggregate change in labor supply at the intensive margin can be decomposed into substitution and income effects, see Section 4.3.1. The substitution effect is estimated at approximately NOK 6.1 billion (EUR 0.60 billion; USD 0.64 billion), while the offsetting income effect amounts to around NOK 0.7 billion (EUR 0.07 billion; USD 0.07 billion), resulting in a net reduction in pre-tax income of NOK 4.7 billion (EUR 0.47 billion; USD 0.49 billion).

Table 5.1 **Tax revenue effects of five percentage-points increase in the two top brackets of the bracket tax, million NOK**

	Revenue
Reference: 2022 tax revenues	622,573
Difference to reference, net revenue effects	
Mechanical effect	7,866
Behavioral effects	-3,620
Intensive margin	-2,931
Extensive margin	-689
Taxes	-548
Benefits	-141
Total effect (mechanical and behavioral)	4,246

Notes: Simulations by Norwegian tax-benefit model LOTTE-Skatt, with 2022-data and the tax system for 2022 as the reference. In the baseline, the two top rates are 46.4% and 47.4%, respectively, and increased to 51.4% and 52.4% by the policy change. Behavioral effects are accounted for by a substitution elasticity of 0.15 and an income elasticity of -0.05 at the intensive margin and a response estimate of 0.2 at the extensive margin.

In line with Thoresen et al. (2010), who estimate the degree of self-financing associated with the tax cuts in Norways 2006 tax reform, we compare the initial mechanical revenue estimate with one that accounts for behavioral responses. This is done by employing the formula $\frac{\Delta M_{BH}}{\Delta M_{NB}}$, where ΔM symbolizes change in revenue and subscripts NB and BH refer to the mechanical (non-behavioral) effect and the behavioral effect, respectively. Using the figures reported in Table 5.1, we estimate the total counteracting behavioral effect to be approximately 43%. In Section 5.4, we revisit the implications of the behavioral effects in terms of interpreting our results with respect to the Laffer curve.

In Table 5.2, we show the distributional effects of this tax increase. As expected, the reductions in disposable income are mainly found in decile group 10. Given that we see that behavioral extensive margin effects are found in decile group 10 too, it means that some individuals are predicted to move from decile group 10 to non-participation due to the tax change. Given that the tax increase begins at income levels around NOK 1 million, one might question the plausibility of individuals opting for non-participation in response to this change. However, these movements are few, only around 800. It is also worth noting that our data includes relatively old employees, as the upper threshold is 61 years. Both Alpert and Powell (2020) and (Messacar, 2024) statistically significant and economically meaningful effects of taxes on labor force participation for older workers. For example, Alpert and Powell (2020) find that older workers is especially responsive to incentives to work, both at the intensive and extensive margins.

[28]The median estimate is 811, with 95% bootstrapped confidence intervals of [754, 870].

Table 5.2 Effects on individual disposable income of five percentage points increase in the two top brackets of the bracket tax. In NOK 1,000.

Reference		Change in disposable income			
Decile group	Disposable income	Mechanical	Intensive	Extensive	Total
1	37.0	-0.1	-0.0	0.0	-0.1
2	193.0	-0.0	-0.0	0.0	-0.0
3	258.9	-0.0	-0.0	0.0	-0.0
4	301.9	-0.0	-0.0	0.0	-0.0
5	346.5	-0.0	-0.0	0.0	-0.0
6	392.5	-0.0	-0.0	0.0	-0.0
7	443.4	-0.0	-0.0	0.0	-0.0
8	502.6	-0.0	-0.0	0.0	-0.0
9	591.7	-0.1	-0.1	0.0	-0.2
10	1,026.8	-17.5	-6.1	-1.7	-25.3
All	409.4	-1.8	-0.6	-0.2	-2.6

Notes: Simulations by Norwegian tax-benefit model LOTTE-Skatt, with 2022-data and the tax system for 2022 as the reference. Behavioral effects are incorporated using a substitution elasticity of 0.15 and an income elasticity of -0.05 at the intensive margin, along with a response estimate of 0.2 at the extensive margin.

5.2 Effects of the work deduction

Next, we assess the effect of adding the work deduction to the 2022 tax scheme – a reform designed to incentivize labor market entry. As anticipated, we observe a substantial effect on the extensive margin due to this reform, with more than 7,500 individuals entering the workforce.[29]

Table 5.3 summarizes the effects on tax revenue resulting from the introduction of the deduction. The direct (non-behavioral) revenue loss is substantial, NOK 23.3 billion (EUR 2.3 billion; USD 2.4 billion). Notably, despite the extensive margin effect, behavioral responses further increase the revenue loss. This arises because the revenue loss from intensive margin responses (see Figure 4.3) exceeds the positive extensive margin effectapproximately NOK 1.9 billion (EUR 0.18 billion; USD 0.19 billion) compared with about NOK 1.1 billion (EUR 0.10 billion; USD 0.10 billion). Consequently, the total revenue loss due to behavioral responses is around NOK 0.74 billion (EUR 0.07 billion; USD 0.08 billion). Appendix B presents robustness tests assessing the extent to which revenue effects and transitions between participation and non-participation depend on key assumptions of the empirical approach.

The intensive margin revenue effect reported in Table 5.3 reflects both substitution and income effect. The substitution effect arises from an effective increase in marginal tax rates as the work deduction phases out, while the income effect results from an overall increase in disposable income. In terms of pre-tax labor income, the negative substitution effect reduces income by approximately NOK 3 billion (EUR 0.30 billion; USD 0.31 billion), while the negative income effect accounts for roughly NOK 1.7 billion (EUR 0.17 billion; USD 0.18 billion).

In Table 5.4, we present the distributional effects of introducing the work deduction. We observe that the behavioral effects are overall positive for the lower decile groups, as some individuals in these decile groups enter the labor market. However, when accounting for both the mechanical and behavioral effects, individuals in the higher decile groups benefit most from the introduction of the work deduction.

[29]The median estimate is 7,529, with 95% bootstrapped confidence intervals ranging from [7,364, 7,698].

Table 5.3 **Tax revenue effects of introducing the work deduction, million NOK**

	Revenue
Reference: 2022 tax revenues	622,537
Difference to reference, net revenue effects	
Mechanical effect	-23,306
Behavioral effects	-743
Intensive margin	-1,861
Extensive margin	1,118
Taxes	230
Benefits	888
Total effect (mechanical and behavioral)	-24,049

Notes: Simulations by Norwegian tax-benefit model LOTTE-Skatt, data for 2022. Behavioral effects are accounted for by a substitution elasticity 0.15 and income elasticity of -0.05 at the intensive margin and 0.2 at the extensive margin.

Table 5.4 **Effects on individual disposable income from introduction of the work deduction. In NOK 1,000**

Reference		Change in disposable income			
Decile groups	Disposable income	Mechanical	Intensive	Extensive	Total
1	37.0	0.3	0.1	0.4	0.7
2	193.0	4.0	0.8	0.4	5.1
3	258.9	3.4	0.0	0.4	3.8
4	301.9	4.7	-0.2	0.3	4.8
5	346.5	6.4	-0.4	0.2	6.1
6	392.5	7.9	-0.7	0.1	7.2
7	443.4	8.3	-1.0	0.0	7.3
8	502.6	7.8	-1.1	0.0	6.7
9	591.7	6.6	-1.3	0.0	5.3
10	1,026.8	3.2	-1.2	0.0	2.1
All	409.4	5.2	-0.5	0.2	4.9

Notes: Simulations by Norwegian tax-benefit model LOTTE-Skatt, data for 2022. Behavioral effects are accounted for by a substitution elasticity 0.15 and income elasticity of -0.05 at the intensive margin and 0.2 at the extensive margin.

5.3 Contrasting results to structural labor supply model simulations

In the following, we compare the simulation results obtained from the "external evidence approach" with those produced by a structural labor supply model, employing the labor supply module from the LOTTE-family of models (Jia et al., 2024; Dagsvik et al., 2014). The module is based on a structural discrete choice framework (Dagsvik and Jia, 2016), a well-established methodological approach for modeling labor supply in microsimulation analysis (van Soest, 1995; Creedy and Kalb, 2005). Whereas Thoresen and Vattø (2015) use estimates of the elasticity of taxable income (ETI) to validate a structural labor supply model, our approach takes the opposite perspective: we use simulated labor supply responses to cross-check results from the "external evidence approach".

Although the methodological approaches are expected to yield at least qualitatively similar results, as discussed in the literature review (Section 2.2), they are not expected to produce identical results. Several factors can explain discrepancies. First, the structural model focuses solely on changes in working hours, whereas the "external evidence approach" is based on changes in labor income, meaning that its response estimates capture variations in both wages and working hours.

As such, the responses we have in mind here are closer to the standard ETI estimates of the quasi-experimental literature (Saez et al., 2012; Neisser, 2021).

Second, a comparison of the tax elasticity estimates obtained from external sources with the response estimates from the labor supply model is made complicated by the fact that the main output of the labor supply model is gross wage elasticities. However, as discussed by Creedy and Mok (2019), the two types of elasticities – the ETI and the gross wage elasticity – reflect comparable behavioral responses at the intensive margin.[30]

Nonetheless, we find it informative to compare the results of the external evidence approach with those from a structural labor supply model. The labor supply model produces the following uncompensated intensive margin wage elasticity estimates across population groups: 0.20 for married females, 0.06 for single females, 0.09 for married males, and 0.01 for single males. These estimates are generally lower than the elasticity assumed in the "external evidence approach" (0.1), except in the case of married females. Furthermore, the extensive margin elasticity for females is estimated at 0.14 for married women and 0.01 for single women.[31]

Third, the structural labor supply model would typically capture (more) long-term structural responses, while quasi-experimental evidence usually reflects short-term frictions and other immediate factors. This suggests that larger estimates from a structural model could align with smaller estimates derived from quasi-experimental evidence measured over a shorter time period (Chetty, 2012). Vattø (2020) and Jia and Vattø (2021) use quasi-experimental evidence and a structural labor supply model, respectively, to discuss the time path of behavioral adjustments, and both studies suggest that the full effect is reached after approximately five years.[32] Finally, we note that there may be differences in the samples used for the estimation too.

Table 5.5 presents results for the same two policy reforms previously discussed in the context of the "external evidence approach". We find that the structural labor supply model generates smaller behavioral responses to the tax policy changes than those following from the "external evidence framework". Nonetheless, it is reassuring that both models produce qualitatively similar outcomes. For increases in tax rates at higher income levels, consistent with the results presented in Section 4.2.1, the structural model predicts reductions in both the total expected number of workers (extensive margin) and the average hours worked per worker (intensive margin).

In the case of the work deduction, the labor supply model predicts an increase in labor supply at the extensive margin, offset by a larger decrease at the intensive margin, yielding a net negative effect on tax revenue. The total behavioral effect is estimated to approximately NOK 600 million (EUR 59 million, USD 62 million), which is close to the effects seen for "external evidence approach" (approximately NOK 750 million, see Table 5.3).

5.4 Applying the procedure to obtain optimal top tax rate

The counteracting behavioral effects of the tax increase presented in Section 5.1 connects to the recurring debate on Laffer-curve results, see for example Goolsbee et al. (1999). The estimated self-financing rate of tax changes is influenced by the initial (pre-reform) tax level. When the initial

[30] The labor supply component of the taxable income elasticity closely mirrors the gross wage elasticity of hours worked (Creedy and Mok, 2019).

[31] Then, overall average elasticities, for the intensive and extensive margins combined, are 0.33 for married females, 0.07 for single females, 0.09 for married males, and 0.01 for single males.

[32] The distinction between short-run and long-run elasticities is also discussed by Kleven et al. (2025). They argue that the returns to effort are delayed and mediated by job switches, such as promotions within firms or movements between firms, which means that short-run micro elasticities (which are the focus of the present study) are different to the true long-run macro elasticity.

Table 5.5 **Changes in labor supply behavior and the accompanying changes in tax revenue for two tax policy changes. Results from structural labor supply model simulation (LOTTE-Arbeid)**

	Increased rates, bracket tax	Work deduction
Change in number of workers (%)	-0.03	0.23
Change in mean hours of work, cond. on working (%)	-0.21	-0.15
Revenue change, behavioral response (million NOK)	-1,860	-596

Notes: Revenue effects of a five percentage-point increase in the two top brackets of the bracket tax and the introduction of the work deduction. Changes in labor supply behavior, measured in %, and effects on tax revenue (million NOK).

tax rate is high, a given rate change generates larger revenue effects, which increases the measured self-financing of the reform. In an extension to this, we may ask how the top rate (fifth step of the bracket tax), both before and after the tax change under consideration here, relates to the peak of the Laffer-curve.

We use the expression for the optimal tax rate of Saez (2001), $\tau^* = \frac{1-g}{1-g+\varepsilon^u+\varepsilon^c(\alpha-1)}$, where g is the welfare weight, α is the Pareto parameter and ε symbolizes the ETI (as before). The Pareto parameter measures the thinness of the top tail, in this case the thinness of the part of the income distribution above the threshold of the fourth bracket of the bracket tax. It is defined as, $\alpha = \frac{z^m}{z^m-\overline{z}}$, where z^m is the average income of the high-income individuals and $\overline{z}$ is the threshold of the fifth step.

Neglecting the welfare weights, or alternatively give no weight to taxpayers at the high end of the income distribution, we derive an expression for the top of the Laffer-curve given by $\tau^* = \frac{1}{1+\varepsilon^u+\varepsilon^c(\alpha-1)}$. When employing the uncompensated elasticity estimate, $\varepsilon^u = \varepsilon^c + \eta' = 0.1$, and the compensated elasticity estimate, $\varepsilon^c = 0.15$, and $\alpha = 2.3$, we get $\tau^* = 0.77$.[33] This implies that the current top marginal tax rate of 47.4% lies well below the peak of the Laffer curve. Even a five percentage-point increase, to 52.4%, would remain clearly below the top of the curve.[34]

6. Conclusion

The starting point of the present study is that policy analysts do not always have access to a fully developed structural simulation model to guide policy-making. Instead of neglecting behavioral effects in the measurement of policies, one may use a procedure we refer to as the "external evidence procedure" for incorporating response estimates in behavioral microsimulation. The phrase reflects that the framework builds on obtaining information on behavioral response from the literature. It is argued that incorporating these externally derived estimates into a non-behavioral tax-benefit model is a practical way for obtaining estimates of effects on tax revenue and income distribution that include behavioral response.

We contend that some form of the "external evidence procedure" is already incorporated into the toolkit of many policy analysts supporting decision-making. However, it likely exists in various forms, including relatively simple back-of-the-envelope calculations. This paper contributes to the literature by providing a detailed exposition of a comprehensive and practical implementation of the approach, highlighting its key inputs.

[33] The estimate of α is obtained from own calculations. Similar estimates from the U.S. are smaller, often around 1.5 (Saez et al., 2012). An alternative is to calculate $\tau^* = \frac{1}{1+\alpha\varepsilon}$, which gives $\tau^* = 0.81$, when ε enters as the uncompensated effect (0.1).

[34] Admittedly, several objections can be raised against this calculation. For example, one may argue that the payroll tax should be included among taxes on labor income.

We have demonstrated the use of the approach by calculating tax revenue effects of extending the Norwegian tax scheme by two policy changes: first, an increase in the two top brackets of the tax on wage income and second, the introduction of a work deduction. In selecting our response estimates, we have referred to previous studies, including those that have estimated intensive margin responses among Norwegian taxpayers, which suggest that responses are moderate. Moreover, we emphasize that calculating counterfactual income of individuals to account for extensive margin effects represents a key challenge of the method. For illustration, we use an extensive margin elasticity of 0.2, and elasticities at the intensive margin of 0.15 for the compensated effect and -0.05 for the income effect.

We then find that the increase in tax rates at the top income levels leads to a counteracting behavioral response, reducing the non-behavioral (mechanical) revenue effect by 43%. In contrast, the behavioral effects of introducing the “work deduction” amplify the negative mechanical effect as the impact of higher marginal tax rates in the phase-out ranges exceeds the effect on the extensive margin. Of course, the results are sensitive to the choice of elasticities and the method of obtaining counterfactual income. Nevertheless, in the case of the work deduction, the extensive margin elasticity would need to be substantially larger than the intensive margin elasticity to generate a positive contribution to tax revenues from behavioral responses. The “external evidence procedure” is useful for specifying these conditions in detail.

In general, we argue that the “external evidence procedure” is a valuable tool for policy analysts involved in policy-making. A key advantage of this approach is its flexibility, as it allows the incorporation of alternative elasticity estimates informed by new research. Furthermore, the method can, in principle, be extended to other behavioral margins, such as income shifting or tax evasion.

Bibliography

Aaberge, R., J. K. Dagsvik, and S. Strøm (1995): "Labor supply responses and welfare effects of tax reforms," *The Scandinavian Journal of Economics*, 97, 635–659.

Aarbu, K. O. and T. O. Thoresen (2001): "Income responses to tax changes evidence from the Norwegian tax reform," *National Tax Journal*, 54, 319–335.

Alpert, A. and D. Powell (2020): "Estimating intensive and extensive tax responsiveness," *Economic Inquiry*, 58, 1855–1873.

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

Auerbach, A. J. (2005): "Dynamic scoring: An introduction to the issues," *American Economic Review*, 95, 421–425.

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

Bartels, C. and N. Pestel (2016): "Short-and long-term participation tax rates and their impact on labor supply," *International Tax and Public Finance*, 23, 1126–1159.

Bastani, S., Y. Moberg, and H. Selin (2021): "The anatomy of the extensive margin labor supply response," *The Scandinavian Journal of Economics*, 123, 33–59.

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

Blau, F. D. and L. M. Kahn (2007): "Changes in the labor supply behavior of married women: 1980–2000," *Journal of Labor Economics*, 25, 393–438.

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

Blundell, R., A. Bozio, and G. Laroque (2013): "Extensive and intensive margins of labour supply: Work and working hours in the US, the UK and France," *Fiscal Studies*, 34, 1–29.

Bourguignon, F. and A. Spadaro (2006): "Microsimulation as a tool for evaluating redistribution policies," *The Journal of Economic Inequality*, 4, 77–106.

Brys, B., S. Matthews, and J. Owens (2011): "Tax reform trends in OECD countries," OECD, Paris.

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

Burtless, G. and J. A. Hausman (1978): "The effect of taxation on labor supply: Evaluating the Gary negative income tax experiment," *Journal of Political Economy*, 86, 1103–1130.

Carroll, R. and W. Hrung (2005): "What does the taxable income elasticity say about dynamic responses to tax changes?" *American Economic Review: Papers and Proceedings*, 95, 426–431.

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

——— (2012): "Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply," *Econometrica*, 80, 969–1018.

Chetty, R., J. N. Friedman, T. Olsen, and L. Pistaferri (2011): "Adjustment costs, firm responses, and micro vs. macro labor supply Elasticities: Evidence from Danish Tax Records," *The Quarterly Journal of Economics*, 126, 749–804.

Chetty, R., A. Guren, D. Manoli, and A. Weber (2013): "Does indivisible labor explain the difference between micro and macro elasticities? A meta-analysis of extensive margin elasticities," *NBER Macroeconomics Annual*, 27, 1–56.

Christiansen, V., Z. Jia, and T. O. Thoresen (2022): "Assessing income tax perturbations," *International Tax and Public Finance*, 29, 472–504.

Creedy, J., N. Gemmell, N. Hérault, and P. Mok (2020): "A microsimulation analysis of marginal welfare-improving income tax reforms for New Zealand," *International Tax and Public Finance*, 27, 409434.

Creedy, J. and G. Kalb (2005): "Discrete hours labour supply modelling: Specification, estimation and simulation," *Journal of Economic Surveys*, 19, 697–734.

Creedy, J. and P. Mok (2019): "Labour supply elasticities in New Zealand," *New Zealand Economic Papers*, 53, 125–143.

Dagsvik, J. K. (1994): "Discrete and continuous choice, max-stable processes, and independence from irrelevant attributes," *Econometrica*, 62, 1179–1205.

Dagsvik, J. K. and Z. Jia (2016): "Labor supply as a choice among latent jobs: Unobserved heterogeneity and identification," *Journal of Applied Econometrics*, 31, 487–506.

Dagsvik, J. K., Z. Jia, T. Kornstad, and T. O. Thoresen (2014): "Theoretical and practical arguments for modeling labor supply as a choice among latent jobs," *Journal of Economic Surveys*, 28, 134–151.

Duan, N. (1983): "Smearing Estimate: A Nonparametric Retransformation Method," *Journal of the American Statistical Association*, 78, 605–610.

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

——— (1999): "Tax avoidance and the deadweight loss of the income tax," *Review of Economics and Statistics*, 81, 674–680.

Figari, F., A. Paulus, and H. Sutherland (2015): "Microsimulation and policy analysis," in *Handbook of Income Distribution*, ed. by A. B. Atkinson and F. Bourguignon, Elsevier, vol. 2, 2141–2221.

Finansministeriet and Skatteministeriet (2024): "Regneprincipper på personskatteområdet," København (in Danish).

Gelber, A. M. (2014): "Taxation and the earnings of husbands and wives: Evidence from Sweden," *Review of Economics and Statistics*, 96, 287–305.

Gelber, A. M., D. Jones, and D. W. Sacks (2020): "Estimating adjustment frictions using nonlinear budget sets: Method and evidence from the earnings test," *American Economic Journal: Applied Economics*, 12, 1–31.

Giertz, S. H. (2007): "The Elasticity of Taxable Income over the 1980s and 1990s," *National Tax Journal*, 60, 743–768.

Goolsbee, A., R. E. Hall, and L. F. Katz (1999): "Evidence on the high-income Laffer curve from six decades of tax reform," *Brookings Papers on Economic Activity*, 1999, 1–64.

Gordon, R. H. and J. B. Slemrod (2000): "Are Real Responses to Taxes Simply Income Shifting Between Corporate and Personal Tax Bases?" in *Does Atlas Shrug?: The Economic Consequences of Taxing the Rich*, ed. by J. Slemrod, Harvard: Harvard University Press, 240–280.

Graber, M., M. Mogstad, G. Torsvik, and O. L. Vestad (2022): "Behavioural responses to income taxation in Norway," *Working Paper, Memorandum, University of Oslo, Department of Economics*, 04/2022.

Gravelle, J. (2014): *Dynamic Scoring for Tax Legislation: A Review of Models*, Washington, DC: Congressional Research Service.

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

Hansson, Å. (2007): "Taxpayers' responsiveness to tax rate changes and implications for the cost of taxation in Sweden," *International Tax and Public Finance*, 14, 563–582.

Harju, J. and T. Matikka (2016): "The elasticity of taxable income and income-shifting: What is real and what is not?" *International Tax and Public Finance*, 23, 640–669.

Hausman, J. A. (1981): "Labor supply," in *How Taxes Affect Economic Behavior*, ed. by H. J. Aaron and J. A. Pechman, Brookings Institution Press, 27–71.

Heckman, J. J. (2010): "Building bridges between structural and program evaluation approaches to evaluating policy," *Journal of Economic Literature*, 48, 356–98.

Heim, B. T. (2007): "The incredible shrinking elasticities. Married female labor supply, 1978–2002," *Journal of Human Resources*, 42, 881–918.

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

Immervoll, H., H. J. Kleven, C. T. Kreiner, and E. Saez (2007): "Welfare reform in European countries: A microsimulation analysis," *The Economic Journal*, 117, 1–44.

Jakobsen, K. M. and J. E. Søgaard (2022): "Identifying behavioral responses to tax reforms: New insights and a new approach," *Journal of Public Economics*, 212.

Jia, Z., B. Larsen, B. Lian, R. Nesbakken, O. E. Nygård, T. O. Thoresen, and T. E. Vattø (2024): "The LOTTE system of tax microsimulation models," *International Journal of Microsimulation*, 17, 73–93.

Jia, Z., T. O. Thoresen, and T. E. Vattø (2025): "Explaining the declining labor supply responsiveness of married women," Manuscript, Statistics Norway.

Jia, Z. and T. E. Vattø (2021): "Predicting the path of labor supply responses when state dependence matters," *Labour Economics*, 71.

Joint Committee on Taxation (2015): "Estimating Changes in the Federal Individual Income Tax: Description of the Individual Tax Model (JCX-75-15)," Washington DC.

——— (2023): "Estimating Changes in the Federal Individual Income Tax: Description of the Individual Tax Model for 2023 (JCX-48-23)," Washington DC.

Keane, M. P. (2010): "Structural vs. atheoretic approaches to econometrics," *Journal of Econometrics*, 156, 3–20.

Klemm, A. D., L. Liu, V. Mylonas, and P. Wingender (2018): "Are Elasticities of Taxable Income Rising?" WP/18/132, International Monetary Fund, Washington D.C.

Kleven, H. (2022): "The EITC and the extensive margin: A reappraisal," NBER Working Paper No 26405, National Bureau of Economic Research, Cambridge MA.

Kleven, H., C. Kreiner, K. Larsen, and J. Søgaard (2025): "Micro versus Macro Labor Supply Elasticities: The Role of Dynamic Returns to Effort," *American Economic Review*, 115, 2849–2890.

Kleven, H. J. (2014): "How Can Scandinavians Tax So Much?" *Journal of Economic Perspectives*, 28, 77–98.

——— (2016): "Bunching," *Annual Review of Economics*, 8, 435–464.

——— (2021): "Sufficient statistics revisited," *Annual Review of Economics*, 13, 515–538.

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

Kniesner, T. J. and J. P. Ziliak (2008): "Evidence of Tax-Induced Individual Behavioral Responses," in *Fundamental Tax Reform: Issue, Choices, and Implications*, ed. by J. Diamond and G. Zodrow, MIT Press, 375–411.

Kreiner, C. T., S. Leth-Petersen, and P. E. Skov (2016): "Tax reforms and intertemporal shifting of wage income: Evidence from Danish monthly payroll records," *American Economic Journal: Economic Policy*, 8, 233–57.

LaLonde, R. J. (1986): "Evaluating the econometric evaluations of training programs with experimental data," *American Economic Review*, 76, 604–620.

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

Lindsey, L. B. (1987): "Individual taxpayer response to tax cuts: 1982–1984: With implications for the revenue maximizing tax rate," *Journal of Public Economics*, 33, 173–206.

Löffler, M., A. Peichl, and S. Siegloch (2018): "The sensitivity of structural labor supply estimations to modeling assumptions," IZA DP No. 11425, IZA Institute of Labor Economics.

Lundberg, J. and J. Norell (2020): "Taxes, benefits and labour force participation: A survey of the quasi-experimental literature," *Journal of the Finnish Economic Association*, 1, 60–77.

Mankiw, N. G. and M. Weinzierl (2006): "Dynamic scoring: A back-of-the-envelope guide," *Journal of Public Economics*, 90, 1415–1433.

Matikka, T. (2018): "Elasticity of taxable income: Evidence from changes in municipal income tax rates in Finland," *The Scandinavian Journal of Economics*, 120, 943–973.

Mertens, K. and J. L. Montiel Olea (2018): "Marginal Tax Rates and Income: New Time Series Evidence," *The Quarterly Journal of Economics*, 133, 1803–1884.

Messacar, D. (2024): "Labor Supply Responses to Income Taxation among Older Couples: Evidence from a Canadian Reform," *American Economic Journal: Economic Policy*, 16, 228–258.

Miao, D., H. Selin, and M. Söderström (2025): "Earnings responses to even higher taxes," *The Economic Journal*, 135, 838–860.

Ministry of Finance (2022): "Et helhetlig skattesystem," NOU 2022:20 (in Norwegian).

Narazani, E., U. Colombino, and B. Palma (2021): "EUROLAB: A multidimensional labour supply-demand model for EU countries," Tech. rep., JRC Working Papers on Taxation and Structural Reforms.

Neisser, C. (2021): "The Elasticity of Taxable Income: A Meta-Regression Analysis," *The Economic Journal*, 131(640), 3365–3391.

Ollonqvist, J., J. Tervola, J. Pirttilä, and T. O. Thoresen (2024): "Accounting for behavioral effects in microsimulation: A reduced form approach," 17, 44–72.

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

Pissarides, C. A. and G. Weber (1989): "An expenditure-based estimate of Britain's black economy," *Journal of Public Economics*, 39, 17–32.

Prescott, E. (2004): "Why do Americans work so much more than Europeans?" *Federal Reserve Bank of Minneapolis Quarterly Review*, 28, 2–13.

Rogerson, R. and J. Wallenius (2009): "Micro and macro elasticities in a life cycle model with taxes," *Journal of Economic Theory*, 144, 2277–2292.

Rubolino, E. and D. Waldenström (2019): "Trends and gradients in top tax elasticities: Cross-country evidence, 1900–2014," *International Tax and Public Finance*, 26, 457–485.

——— (2020): "Tax progressivity and top incomes evidence from tax reforms," *The Journal of Economic Inequality*, 18, 261–289.

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

——— (2002): "Optimal income transfer programs: Intensive versus extensive labor supply responses," *The Quarterly Journal of Economics*, 117, 1039–1073.

——— (2010): "Do taxpayers bunch at kink points?" *American Economic Journal: Economic Policy*, 2, 180–212.

Saez, E., J. Slemrod, and S. H. Giertz (2012): "The elasticity of taxable income with respect to marginal tax Rates: A critical review," *Journal of Economic Literature*, 50, 3–50.

Slemrod, J. (1996): "High-income families and the tax changes of the 1980s: The anatomy of behavioral response," in *Empirical Foundations of Household Taxation*, ed. by M. Feldstein and J. Poterba, University of Chicago Press, 169–189.

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

Statistics Norway (2019): *Income and wealth statistics for households*, https://www.ssb.no/en/inntekt-og-forbruk/statistikker/ifhus.

Thoresen, T. O. (2004): "Reduced tax progressivity in Norway in the nineties: The effect from tax changes," *International Tax and Public Finance*, 11, 487–506.

Thoresen, T. O., J. Aasness, and Z. Jia (2010): "The short-term ratio of self-financing of tax cuts: An estimate for Norway's 2006 tax reform," *National Tax Journal*, 63, 93–120.

Thoresen, T. O., E. E. Bø, E. Fjærli, and E. Halvorsen (2012): "A suggestion for evaluating the redistributional effects of tax changes: With an application to the 2006 Norwegian tax reform," *Public Finance Review*, 40, 303–338.

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

van Soest, A. (1995): "Structural models of family labor supply: A discrete choice approach," *Journal of Human Resources*, 30, 63–88.

Vattø, T. E. (2020): "Estimating the elasticity of taxable income when earnings responses are sluggish," *FinanzArchiv/Public Finance Analysis*, 76, 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.

Whalen, C. and F. Reichling (2017): "Estimates of the Frisch elasticity of labor supply: A review," *Eastern Economic Journal*, 43, 37–42.

Appendix

A. More details on the empirical approach to obtain extensive margin effects

A.1 Summary statistics

In this appendix, we provide further details on how we obtain estimates of the effects on the extensive margin, i.e., how we impute incomes in the unobserved state of the individuals – obtaining measures for individual "counterfactual income".

We begin by presenting data characteristics for the categories of individuals used in the procedure for extensive margin effects. As outlined in Section 4.4, we impute incomes for individuals in unobserved states using the estimation results and supplementary empirical evidence, drawing on data for individuals who transitioned between participation and non-participation between 2019 and 2022. Table A1 reports summary statistics for the four groups in the data. Two of the groups are samples used for estimation: individuals who moved from non-participation in 2019 to participation in 2022 ($NP \rightarrow P$) and those who transitioned from participation in 2019 to non-participation in 2022 ($P \rightarrow NP$). The two other groups consist of individuals for which econometric estimates are used to impute income: labor earnings for individuals not in the labor market (NP) and income in non-participation for participants (P) in 2022. As explained in Section 4.4, we exploit the panel structure of the data, i.e., that individuals are observed in both 2019 and 2022, to obtain econometric evidence, which is then used to assign counterfactual incomes to individuals in the tax-benefit model dataset.

Table A1 presents both data characteristics and prediction results for the four groups. First, when comparing the average labor earnings of groups P and ($NP \rightarrow P$), we observe differences in earnings levels. As expected, individuals who have shifted into participation have lower wages – we find a 36% difference between the groups. We cannot determine whether this difference arises from hours worked or hourly wages. However, it is worth noting that the ($NP \rightarrow P$) subsample receives substantially higher disability benefit and social assistance than the group P,[35] is twice as likely to have an immigration background, and has, on average, roughly one year less schooling.

Next, we draw attention to the fact that the predicted and observed labor earnings for the ($NP \rightarrow P$) group are very close. This indicates that the prediction procedure relatively precisely captures the in-sample wage distribution. Using the model out-of-sample to predict wage income for the NP group reveals a 10% difference in average predicted earnings between the ($NP \rightarrow P$) and NP groups, indicating that these populations differ on observables. One could argue that positive selection on unobservables occurs in the estimation sample, biasing predicted wage estimates for the NP sample upward. However, an average predicted wage of around NOK 346,000 is quite low in the Norwegian context. Compared to the ($NP \rightarrow P$) sample, the NP sample is about four years older and has two years less education, on average. Additionally, the estimation sample ($NP \rightarrow P$) is 8 percentage points more likely to include individuals with an immigration background, compared to NP.

The ($P \rightarrow NP$) sample is used to predict the probability of receiving disability benefit for people leaving the labor market. Compared to individuals in the NP group, they are 10 percentage

[35] Recall that social assistance refers to the combined amount of direct financial assistance from the Norwegian Labour and Welfare Administration (økonomisk sosialhjelp) and housing allowance (bostøtte) from The Norwegian State Housing Bank (Husbanken).

points less likely to receive disability benefit, 6 percentage points less likely to receive social assistance, and have 1.2 years more schooling on average. The estimated in-sample probability of disability take-up is 0.47 on average. Using the model to predict out-of-sample for the P group (i.e., in the hypothetical case of not participating for those observed participating in the labor market), we obtain a 6 percentage-point lower probability of receiving disability benefit. The P dataset is characterized by a greater share of males, a lower likelihood of individuals having an immigration background, and, on average, one extra year of education. – all factors that, according to the logistic model, are associated with a lower probability of receiving disability benefit; see Table A2.

A.2 Regression results

A.2.1 Wage regression

Table A2 shows estimation results for regressions used to predict counterfactual labor income and probabilities of receiving disability benefit. The dataset ($NP \rightarrow P$) is used to fit a wage model and the dataset ($P \rightarrow NP$) is used to fit a model for disability benefit model take-up. Except for potential work experience and previous wage income, all other variables are dummy variables.

We begin with the model used to assign wage income to non-participants. The log wage income model is a log-linear specification where a one-unit increase in X_i is associated with a $(e^{\beta_i} - 1) \times 100\%$ change in labor income. For small coefficients ($\beta < 0.1$), this approximates the percentage change in labor earnings.

Table A2 shows that work experience follows a typical life-cycle pattern: each additional year is associated with a 1.1% increase in earnings, which declines by 0.02% with the square of experience, indicating diminishing returns and peaking at around 28 years of potential experience. For field of study, the reference category is "Unspecified field of study". "Humanities" is associated with 5.6% lower labor earnings, "Primary industries" shows no significant wage differential, and "Other fields" shows positive wage premiums ranging from 5.4% ("General subjects") to 19% ("Transport/communication").

For education level, the reference category combines no education and pre-school education. There are no significant wage differences between the reference category and primary, lower secondary, upper secondary basic, and upper secondary final education. Post-secondary non-tertiary education is associated with a 11% wage increase. Higher education shows substantial premiums: undergraduate degrees (15%), graduate degrees (30%), and postgraduate qualifications (35%) all yield significant wage increases. Unspecified education level is associated with a 17% wage premium.

Females have 18% lower wage income than men, on average, and having received disability benefit in 2019 gives 13–51% lower wage income compared to people receiving no disability benefit in 2019. Compared to living in the most central areas, less central areas implies a lower wage, between 5% and 10%, depending on a centrality index. Compared with being Norwegian-born to two Norwegian parents, having an immigration background is associated with a 4–11% lower wage income. This effect does not apply to individuals who are Norwegian-born with one foreign parent.

A.2.2 Logistic regression for predicting probability of receiving disability benefit

Next, we discuss results for the model used to assign disability benefit to individuals not participating in the labor market. This is done by a logistic regression, where the dependent variable is a binary variable that takes the value 1 if the disability benefit is larger than zero in 2022, and 0 otherwise. The estimation results are reported in Table A2. The reported coefficients are raw es-

Table A1 **Summary statistics for estimation and prediction samples, 2022**

	NP P	NP	P NP	P
log(labor earnings +1)	12.73	2.88	5.42	13.27
	(0.55)	(4.62)	(5.20)	(0.49)
log(predicted labor earnings)	12.74	12.62	–	–
	(0.53)	(0.53)		
Labor earnings (NOK 1,000)	392.69	11.69	24.03	656.17
	(221.58)	(25.18)	(33.12)	(413.81)
Predicted labor earnings (NOK 1,000)	391.84	345.63	–	–
	(214.21)	(191.06)		
Female	0.51	0.53	0.49	0.47
	(0.50)	(0.50)	(0.50)	(0.50)
Education (years)	12.32	10.62	11.86	13.46
	(4.49)	(4.68)	(4.22)	(4.13)
Disability benefit dummy	0.12	0.57	0.47	0.04
	(0.33)	(0.50)	(0.50)	(0.19)
Disability benefit (NOK 1,000)	21.25	170.43	135.91	6.65
	(65.16)	(162.06)	(160.50)	(37.32)
Prob(disability benefit)	–	–	0.47	0.41
			(0.29)	(0.23)
Social assistance (dummy)	0.11	0.19	0.13	0.05
	(0.31)	(0.39)	(0.33)	(0.21)
Social assistance (NOK 1,000)	4.10	12.61	6.17	2.47
	(20.26)	(44.96)	(27.11)	(14.11)
Immigration background	0.47	0.39	0.41	0.27
	(0.50)	(0.49)	(0.49)	(0.44)
Age	39.45	44.60	43.66	43.08
	(9.92)	(10.75)	(10.43)	(10.30)
Number of observations	99,993	539,529	106,497	2,019,756

Notes: P = Participants in the labor market, NP = non-participants in the labor market. (NP P) = observed participants who transitioned from non-participation in 2019 to participation in 2022. (P NP) = observed non-participants who transitioned from participation to non-participation in the period 2019–2022. Disability benefit is the observed sum of disability pension and work assessment allowance. Social assistance is the observed sum of financial assistance and housing allowance. Prob(disability benefit) is the predicted probability of receiving disability benefit in the NP alternative, estimated by logistic regression in Table A2. Dummy indicates a binary variable equal to 1 when the underlying variable is greater than 0. Immigration background equals 0 for individuals who are Norwegian-born to Norwegian-born parents. – indicates that the variable is not available for this group. Standard deviations in parentheses.

timates and represent the change in log-odds for a one-unit change in the predictor variable. Positive/negative coefficients indicate an increased/decreased probability of receiving disability benefit.

The probability of receiving benefit is increasing at a decreasing rate over the full support of years of experience. Compared to "Unspecified field of study" all other fields significantly increase the probability of receiving disability benefit. The exceptions are "Primary industries", which lowers the probability, and "Transport/communication", which has no significant effect on the probability of receiving the benefit.

Except for primary and lower secondary education, all other education levels decrease the probability of receiving disability benefit. Being female is associated with an increased probability, as are all centrality levels (compared to the most central areas). The association with previous wage income is increasing and concave. The marginal effect is positive up to a wage income of around NOK 400,000, after which the marginal effect turns negative.

Being an immigrant, foreign-born to one Norwegian parent, or Norwegian-born to immigrants all significantly reduce the probability of receiving disability benefit, compared to Norwegian-born to Norwegian parents. Finally, having previously received disability benefit is a strong and significant predictor of receiving disability benefit in 2022.

Table A2 **Regression estimation results used to impute counterfactual wage and probability of receiving disability benefits**

	Log labor income (2022) *OLS*		Disability benefit (2022) *Logistic*	
	Coef.	(SE)	Coef.	(SE)
Work Experience				
Potential experience (years)	0.011***	(0.001)	0.064***	(0.004)
Potential experience (years)^2	-0.000***	(0.000)	-0.000***	(0.000)
Field of Study				
General subjects	0.054***	(0.009)	0.265***	(0.042)
Humanities and arts	-0.056***	(0.010)	0.218***	(0.051)
Teacher training	0.070***	(0.011)	1.051***	(0.055)
Social sciences and law	0.098***	(0.011)	0.360***	(0.061)
Business and administration	0.137***	(0.009)	0.213***	(0.046)
Natural sciences/technical	0.191***	(0.008)	0.407***	(0.041)
Health, welfare and sport	0.157***	(0.009)	0.953***	(0.044)
Primary industries	0.011	(0.016)	-0.138**	(0.070)
Transport/communications	0.195***	(0.012)	0.056	(0.052)
Education Level				
Primary	-0.030	(0.031)	-0.188	(0.141)
Lower secondary	-0.019	(0.028)	-0.245	(0.126)
Upper secondary (basic)	-0.058	(0.030)	-0.574***	(0.131)
Upper secondary (final)	0.008	(0.028)	-0.550***	(0.124)
Post-secondary non-tertiary	0.107***	(0.029)	-0.669***	(0.131)
Tertiary (undergraduate)	0.146***	(0.028)	-0.911***	(0.126)
Tertiary (graduate)	0.301***	(0.028)	-1.705***	(0.129)
Postgraduate	0.354***	(0.033)	-1.779***	(0.168)
Unspecified	0.165***	(0.028)	-1.448***	(0.127)
Demographics				
Female	-0.178***	(0.003)	0.665***	(0.017)
Previous Disability Benefit (2019)				
0-0.5G	-0.131***	(0.017)	2.394***	(0.079)
0.5G-1G	-0.147***	(0.016)	2.648***	(0.084)
1G-2G	-0.193***	(0.007)	3.670***	(0.074)
2G-3G	-0.242***	(0.007)	4.500***	(0.132)
3G-4G	-0.195***	(0.010)	5.075***	(0.321)
4G+	-0.510***	(0.058)	4.736***	(1.032)
Geographic Controls				
Centrality: 02	-0.046***	(0.005)	0.239***	(0.023)
Centrality: 03	-0.066***	(0.005)	0.452***	(0.023)
Centrality: 04	-0.084***	(0.006)	0.476***	(0.026)
Centrality: 05	-0.091***	(0.007)	0.364***	(0.031)
Centrality: 06	-0.097***	(0.009)	0.228***	(0.040)
Immigration Status				
Immigrants	-0.112***	(0.004)	-0.723***	(0.019)
Norwegian-born to immigrants	-0.044***	(0.011)	-0.204***	(0.066)
Foreign-born, one Norwegian parent	-0.075***	(0.020)	-0.284***	(0.092)
Norwegian-born, one foreign parent	-0.008	(0.009)	-0.011	(0.041)
Foreign-born to Norwegians	-0.047***	(0.018)	-0.110	(0.080)
Previous Wage Income (2019)				
Labor income	—	—	15.898***	(0.564)
Labor income squared	—	—	-0.585***	(0.021)
Constant	12.682***	(0.028)	-109.632***	(3.765)
Model Statistics				
Observations	99,993		106,497	
Adjusted Rš	0.128		—	
Pseudo Rš	—		0.284	
Standard errors	Robust		IID	

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Dependent variables: OLS model uses log of annual labor earnings in 2022; logistic model takes value 1 for receipt of disability pension (Uføretrygd) or work assessment allowance (Arbeidsavklaringspenger) in 2022. Potential work experience = max(0, age - max(18, years of education + 6)). Reference categories: Field – Unspecified field of study; Education – No education and pre-school education; Gender – Male; Disability benefits – 0; Centrality – 01, Centrality 01 (highest), Centrality 06 (lowest); Immigration – Norwegian-born to Norwegian parents. Wage income variables use inverse hyperbolic sine (asinh) transformation of 2019 wage income. Disability benefits and wage income reflect 2019 values.

A.3 Social assistance imputation

We next describe how social assistance is imputed in our analysis. Social assistance refers to the combined financial support from the Norwegian Labour and Welfare Administration and the housing allowance provided by The Norwegian State Housing Bank (Husbanken). Both support schemes are means-tested and administered through individual case processing, making them difficult to model explicitly in a microsimulation context.

Social assistance appears in two different components of our empirical approach. First, as explained in Section 4.4, if individuals in the participation state are not assigned disability benefit when not participating, some receive social assistance. Second, when individuals in non-participation are assigned a wage income in the participation state, we adjust (i.e., scale down) the social assistance in participation based on their observed social assistance in non-participation.

To estimate the amount of social assistance received by individuals who leave employment without transitioning to disability benefits, we further restrict the $(P \rightarrow NP)$ sample by excluding individuals who received disability benefit in 2022. To estimate the probability and amount of social assistance among the working-age population, we employ a two-step correction function approach. In the first step, a logistic regression predicts the likelihood of receiving social assistance. In the second step, we estimate the logarithm of the social assistance amount, conditional on receipt, using an OLS model that includes the predicted probability from the first stage to account for potential correlation between the likelihood of receiving benefits and the amount received. Table A3 reports estimation results with demographic and geographic controls, while Table A4 (continuing Table A3) presents estimates related to household composition and income. Standard errors in the second step are bootstrapped with 1,000 replications to account for uncertainty in the two-step procedure.

When using the model to impute counterfactual social assistance, we perform Bernoulli trials on the predicted probability of receiving the benefit for each individual; see Section 4.4. The consequence of this is that a group is established, consisting of individuals who receive neither disability benefits nor social assistance in the non-participation state. We predict social assistance at the log-level, and use a Duan factor (Duan, 1983) of 1.8 when re-transforming to levels.

We also examine the reduction in social assistance when individuals transition from non-participation to employment. This is done for a subset of the $(NP \rightarrow P)$ group – those who received social assistance in 2019 and who are observed working in 2022. Columns 1 and 2 of Table A5 report the average social assistance (in 1,000 NOK) for this group while employed, broken down by centrality level.[36] Column 3 of Table A5 shows the factors (based on the 2022-to-2019 ratio) for which we use to downscale social assistance in the state of working: amounts are multiplied by the downscaling factor. For individuals who did not receive any assistance, the values are still zero. For recipients, the imputed amount is adjusted downward by 62–72%.

[36] In practice, social assistance is determined by municipal schemes rather than by the degree of centrality. However, due to data limitations, we use centrality as a proxy. For instance, housing allowances are directed toward individuals with low income and high housing costs, the latter generally varying by centrality.

Table A3 **Regression estimation results used to impute probability of receiving financial assistance and assistance amounts (Part A: Demographics and geographic controls).**

	Financial assistance receipt *Logistic*		Log financial assistance amount *OLS*	
	Coef.	(SE)	Coef.	(SE)
Demographics				
Age (linear)	-0.336***	(0.110)	0.097***	(0.018)
Age (squared)	0.008***	(0.003)	-0.001***	(0.000)
Age (cubed)	-0.000***	(0.000)	—	—
Female	-0.116***	(0.040)	-0.069	(0.040)
Field of Study				
General subjects	0.238***	(0.072)	—	—
Humanities and arts	0.212***	(0.095)	—	—
Teacher training	-0.053	(0.140)	—	—
Social sciences and law	-0.018	(0.125)	—	—
Business and administration	0.251***	(0.086)	—	—
Natural sciences/technical	0.007	(0.073)	—	—
Health, welfare and sport	0.059	(0.092)	—	—
Primary industries	-0.880***	(0.174)	—	—
Transport/communications	-0.129	(0.109)	—	—
Education Level				
Primary	-0.241	(0.240)	—	—
Lower secondary	-0.405	(0.217)	—	—
Upper secondary (basic)	-0.910***	(0.240)	—	—
Upper secondary (final)	-0.852***	(0.213)	—	—
Post-secondary non-tertiary	-1.049***	(0.237)	—	—
Tertiary (undergraduate)	-1.186***	(0.217)	—	—
Tertiary (graduate)	-1.665***	(0.225)	—	—
Postgraduate	-1.657***	(0.301)	—	—
Unspecified	-1.971***	(0.221)	—	—
Geographic Controls				
Centrality: 02	0.091**	(0.044)	-0.170***	(0.043)
Centrality: 03	0.010	(0.046)	-0.303***	(0.044)
Centrality: 04	-0.194***	(0.057)	-0.403***	(0.058)
Centrality: 05	-0.261***	(0.069)	-0.428***	(0.070)
Centrality: 06	-0.479***	(0.091)	-0.406***	(0.102)
Immigration Status				
Immigrants	-0.621***	(0.039)	—	—
Norwegian-born to immigrants	-0.201	(0.116)	—	—
Foreign-born, one Norwegian parent	-0.099	(0.184)	—	—
Norwegian-born, one foreign parent	0.154	(0.083)	—	—
Foreign-born to Norwegians	0.104	(0.166)	—	—

Notes: Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Table continues on next page with household composition and income variables.

Table A4 **Regression estimation results used to impute probability of receiving financial assistance and assistance amounts (Part B: Household composition and income controls).**

	Financial assistance receipt		Log financial assistance amount	
	Logistic		*OLS*	
	Coef.	(SE)	Coef.	(SE)
Household Composition				
Living alone, 30-44 years	0.005	(0.087)	-0.116	(0.083)
Living alone, 45-66 years	0.005	(0.098)	-0.267***	(0.100)
Couple no children, <30 years	-0.097	(0.151)	-0.068	(0.151)
Couple no children, 30-44 years	-0.157	(0.113)	-0.147	(0.103)
Couple no children, 45-66 years	-0.895***	(0.129)	-0.146	(0.138)
Couple no children, 67+ years	-1.555***	(0.411)	-0.013	(0.684)
Married couple, small children	-0.568***	(0.139)	-0.131	(0.107)
Cohabiting couple, small children	-0.313**	(0.132)	-0.195**	(0.089)
Married couple, older children	-0.756***	(0.125)	-0.070	(0.119)
Cohabiting couple, older children	-0.238	(0.131)	-0.144	(0.118)
Lone mother, small children	1.204***	(0.140)	-0.411***	(0.103)
Lone father, small children	0.589**	(0.284)	-0.091	(0.293)
Lone mother, older children	1.155***	(0.124)	-0.101	(0.108)
Lone father, older children	0.657***	(0.165)	-0.072	(0.151)
Married couple, adult children	-0.659***	(0.123)	-0.416***	(0.139)
Cohabiting couple, adult children	-0.058	(0.167)	-0.551***	(0.168)
Lone mother, adult children	-0.084	(0.119)	-0.376***	(0.117)
Lone father, adult children	-0.147	(0.168)	-0.344*	(0.196)
Multi-person families	-0.720***	(0.121)	-0.153	(0.130)
Other multi-family, no children	-0.701***	(0.151)	-0.853***	(0.185)
Multi-family, small children	-0.596***	(0.152)	-0.183	(0.150)
Multi-family, older children	-0.487***	(0.163)	-0.312*	(0.188)
Income and Benefits				
Log labor income (lag)	-0.372***	(0.031)	-0.180***	(0.037)
Children under 6	0.147**	(0.062)	—	—
Children under 18	0.205***	(0.031)	—	—
Partner wage (linear)	-0.907***	(0.151)	—	—
Partner wage (squared)	0.169***	(0.025)	—	—
Partner wage (cubed)	-0.008***	(0.001)	—	—
Social assistance (lag)	0.162***	(0.005)	—	—
Housing support (lag)	0.157***	(0.006)	—	—
Primary market income (linear)	-2.745**	(1.106)	—	—
Primary market income (squared)	0.378***	(0.144)	—	—
Primary market income (cubed)	-0.013***	(0.005)	—	—
Financial Assistance Model				
Probability of assistance (linear)	—	—	4.082***	(0.570)
Probability of assistance (squared)	—	—	-6.851***	(1.351)
Probability of assistance (cubed)	—	—	4.555***	(0.900)
Constant	8.819***	(1.515)	10.266***	(0.583)
Model Statistics				
Observations	55,991		6,421	
Adjusted Rš	—		0.145	
Pseudo Rš	0.278		—	
RMSE	—		1.250	
Standard errors	IID		Bootstrap	

Notes: Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Dependent variables: Logistic model indicates receipt of financial assistance (fin_ass_dummy) during the period; OLS model uses log of financial assistance amount for recipients only. Age variables use polynomial terms. Partner wage and primary market income variables use inverse hyperbolic sine (asinh) transformation. Reference categories: Household – Living alone, person 67 years and over; Education – No education and pre-school education; Field – Unspecified field of study; Gender – Male; Centrality – 01 (highest); Immigration – Norwegian-born to Norwegian parents.

Table A5 **Social assistance in 2019 and 2022 (in NOK 1,000) by centrality level**

	$(NP \to P \mid \mathrm{SA}_{2019} > 0)$		
Centrality	2019	2022	Downscale factor
01 (highest)	75.69	27.49	0.36
02	66.22	20.15	0.30
03	58.05	16.96	0.29
04	53.43	13.27	0.25
05	51.31	15.53	0.30
06 (lowest)	49.50	12.58	0.25

Notes: Social assistance in 2019 and 2022 are measured in NOK 1,000. Downscaling factors obtained from a subset of individuals: observed working in 2022 and not working in 2019, in combination with receiving social assistance in 2019 ($NP \to P \mid \mathrm{SA}_{2019} > 0$), differentiated by centrality index (Centrality codes: 01 = most central, 06 = most rural). The downscaling factor is calculated as the ratio of social assistance levels in 2022 to those in 2019. Social assistance for the in-work state is obtained by multiplying the observed social assistance in non-participation by this factor. SA = social assistance.

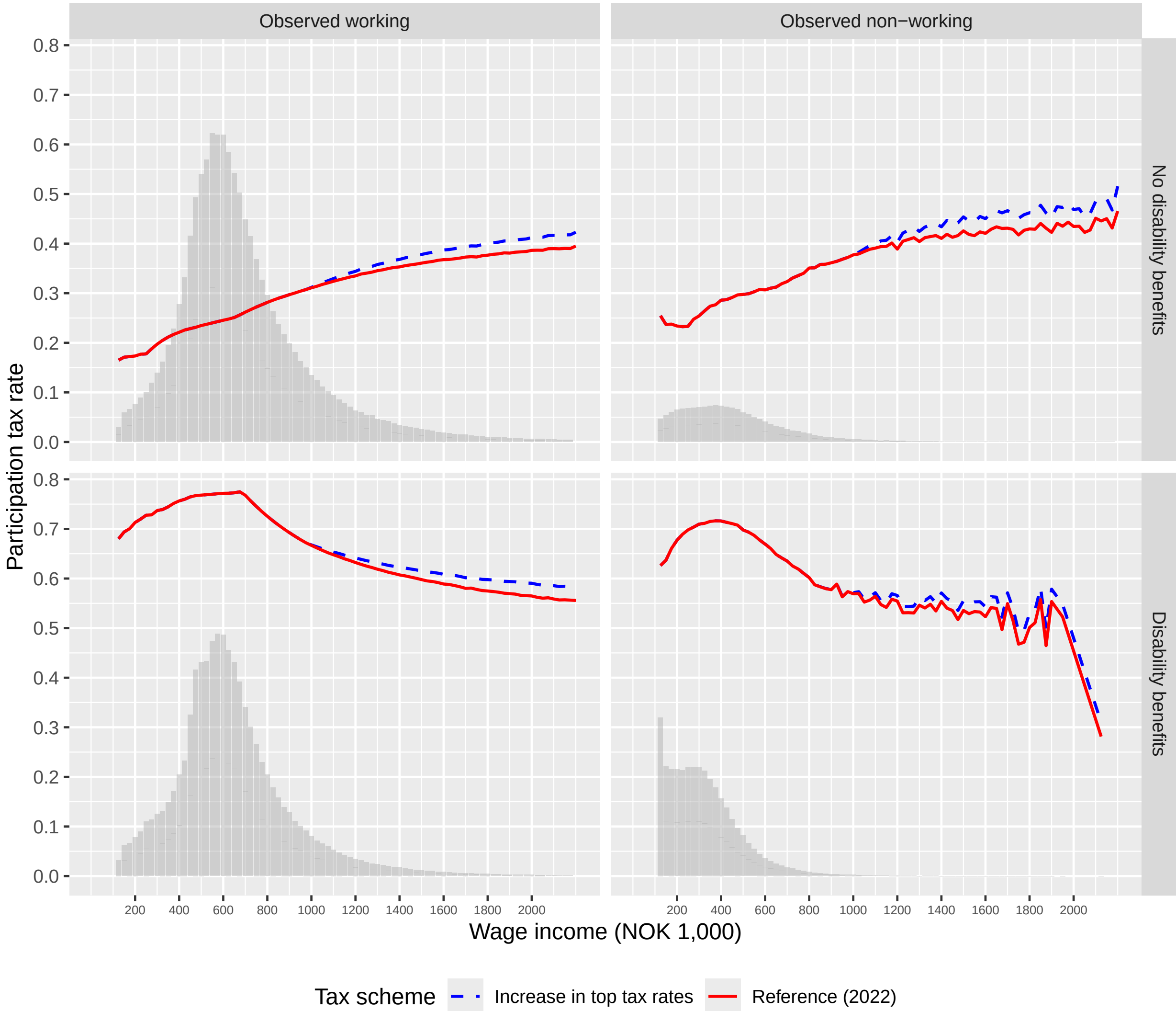
A.4 Participation tax rates for different groups

Given the central role of participation tax rates in our empirical approach for the extensive margin calculations, Figures A1 and A2 provide detailed information on participation tax rates under different tax schemes, comparing the benchmark scheme with the two tax policy changes used for illustration. Each figure presents results for four groups, defined by labor market participation and whether individuals receive disability benefits. The horizontal axis refers to observed income for the working population and imputed wage income for the non-working population.

For the working population, disability benefit status is simulated via a Bernoulli trial using estimated probabilities from the logistic model reported in Table A2; see details in Section 4.4. For the non-working population, status is based on observed disability benefit receipt.

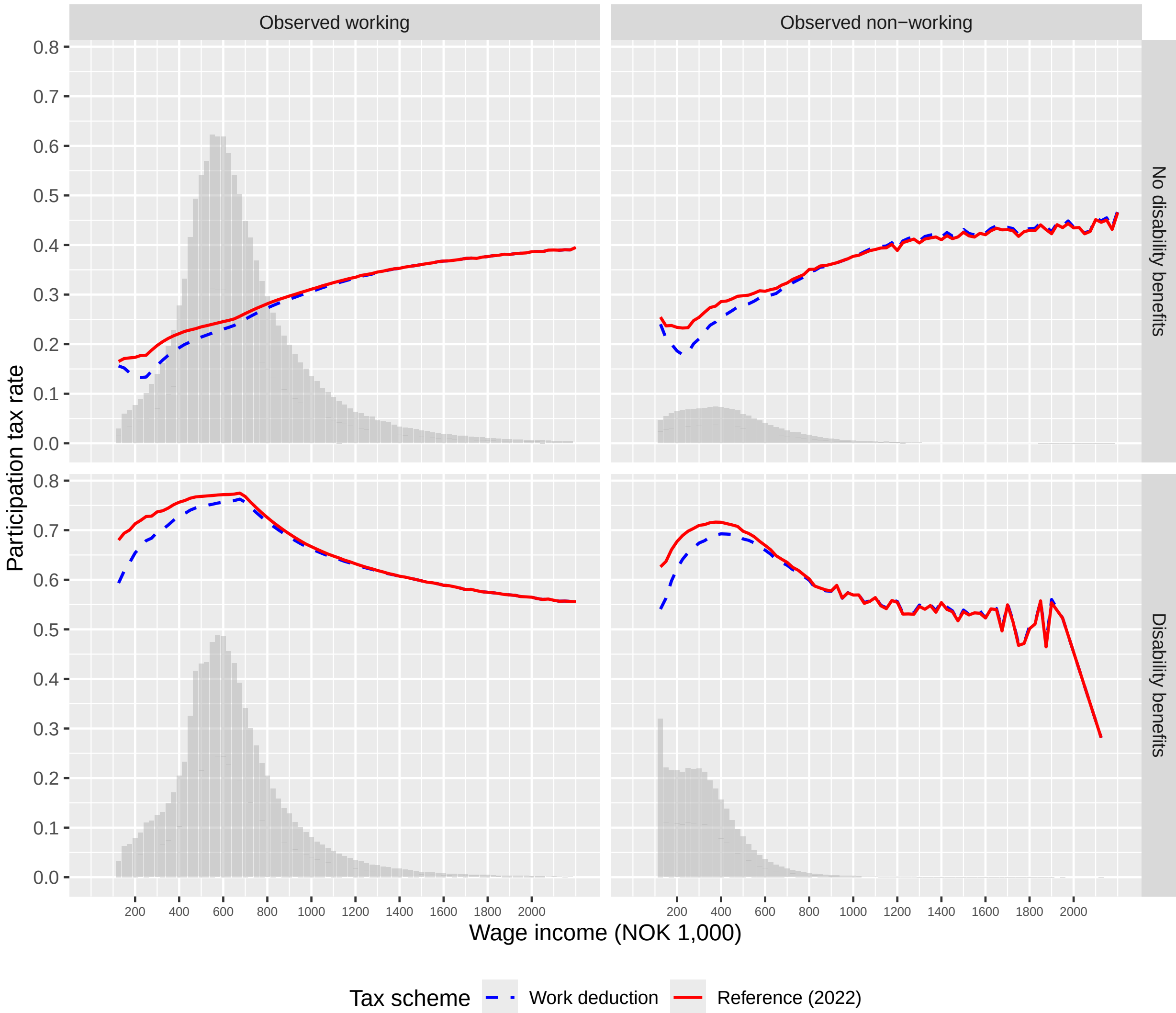
The figures describe how the alternative tax schemes affect the PTRs. When non-participation excludes disability benefit, PTRs rise with the average tax rate for both working and non-working groups. However, the observed non-working population starts from a slightly higher PTR level, creating a small upward shift. When non-participation includes disability benefit, PTRs start at a much higher level due to the phase-out of the benefit. PTRs increase with the average tax rate until incomes reach about NOK 600,000, at which point the disability benefit is fully phased out. This produces a sharp drop in the effective marginal tax rate, which in turn lowers the effective average tax rate and generates the subsequent decline in the PTRs.

Figure A1 **Participation tax rate changes for a five percentage-point increase in top tax rates**



Notes: This figure shows participation tax rates (PTR) under the 2022 tax scheme and for the two policy alternatives. For each wage level, the participation tax rate represents the median rate within the group defined by the panel. The blue dashed line shows how the PTR changes when introducing a five percentage-point increase in top tax rates. The horizontal axis displays earned income for the observed working population and imputed earned income for the observed non-working population. For the observed working population, non-participation (disability benefit/no disability benefit) is determined by Bernoulli trials, based on estimated probabilities from the logistic regression in Table A2. The observed non-working population is categorized based on their current disability benefit status. The gray density curve represents the wage distribution.

Figure A2 **Participation tax rate changes for the working tax deduction**



Notes: This figure shows participation tax rates (PTR) under the 2022 and for the two policy alternatives. For each wage level, the participation tax rate represents the median rate within the group defined by the panel. The blue dashed line illustrates how the PTR changes when the work deduction is introduced. The horizontal axis displays earned income for the observed working population and imputed earned income for the observed non-working population. For the observed working population, non-participation (disability benefit/no disability benefit) is determined by Bernoulli trials based on estimated probabilities from the logistic regression in Table A2. The observed non-working population is categorized based on their current disability benefit status. The gray density curve represents the wage distribution.

B. Sensitivity checks

B.1 Sensitivity to the assignment of counterfactual income

The empirical estimates of revenue effects presented in the main analysis are sensitive to several methodological choices, including the elasticity estimates used and the methods for assigning counterfactual income. It is therefore important to evaluate the extent to which the results depend on these choices. In the following, we describe a selection of sensitivity tests conducted to explore this issue.

We first look at the effects of the alternative where the two top rates of the bracket change, using different ways of defining counterfactual income. Table B1 reports results with respect to revenue and number of individuals shifting due to the extensive margin effect under different assumptions about income in the non-participation state. Rather than linking access to disability benefit and social assistance to individual probabilities, we consider a case where all working individuals receive disability benefit when not working. Then, we show results for an alternative where workers (in the counterfactual state of non-participation) get disability benefit with the probability $\widehat{q}_{di}$ and social assistance support with certainty if they do not receive disability benefits. We also present results for a scenario in which non-participants have no access to these support schemes.

Table B1 clearly shows that both revenue and the number of shifts depend on the empirical approach. As expected, more generous non-participation benefits lead to larger government revenue losses. Most notably, when the disability benefit is assigned to all with a potential to exit, the extensive margin revenue loss increases from NOK 689 million (EUR 68 million; USD 72 million) in the baseline to over NOK 1,000 million (EUR 104 million; USD 109 million). When income in non-participation increases, it induces approximately 30% more individuals to exit the work option, with all who leave receiving maximum disability benefit. Decomposing this revenue loss further, reveals that roughly 44% stems from increased disability benefit expenditures, while the remaining 56% results from foregone tax revenues.

The baseline scenario and the social assistance to all scenario yield virtually identical results, with no discernible differences in either revenue effects or switching behavior. In contrast, the none to all scenario, as expected, produces the smallest revenue effects. With no benefit costs in the non-participation state and, correspondingly, few switches due to the reduced attractiveness of leaving work, the overall fiscal impact is minimal.

B.2 Sensitivity to the threshold for participation

Next, we examine the effect of how participation is defined. As described in Section 4.4, we define the participation threshold at 1 G in the reference scenario – recall that G denotes the basic amount of the Norwegian National Insurance Scheme. Table B2 presents the effects of alternative participation definitions on the extensive margin responses to the introduction of the work deduction.

Changing the participation threshold creates two simultaneous effects that affect the revenue and the choice to change state at the extensive margin. First, there is an effect coming from more individuals being classified as non-participating when the threshold increase. This expands the pool of potential switchers and leads to higher switching rates, all else equal. Second, there are individual-level incentive effects operating through two channels. The labor market composition changes as fewer working-age individuals are initially employed (when the threshold for participation increases), altering individual switching incentives through the P term in Equation (3.11). Additionally, the

Table B1 **Revenue effects of five percentage-points increase in the two top brackets of the bracket tax. Effects on revenue and switches for alternative counterfactual income in the state of non-participation**

Scenario	Revenue	Switching individuals
Baseline (disability benefits, social assistance or none)	-689	811
	(-724, -599)	(754, 870)
Disability benefits to all	-1,051	1,062
	(-1,126, -980)	(1,004, 1,126)
Disability benefits or social assistance to all	-681	823
	(-745, -625)	(768, 879)
None to all	-483	689
	(-538, -435)	(644, 738)

Notes: Top row shows median values (50%), bottom row shows confidence intervals (2.5%, 97.5%). 1 G is used as threshold for participation.

higher threshold effectively truncates the bottom of the wage distribution, causing the regression model to predict higher average wages in participation. The alternative wage distribution for 2 G explains that there are fewer individuals changing state in the 2 G alternative than in the benchmark (1 G). Individuals in scenario 2 G enter at higher points in the wage distribution, see last column of Table B2, where the work deduction generates smaller changes in the participation tax rate (PTR), resulting in weaker switching incentives and fewer individuals moving. Figure B1 provides additional detail on the distribution of earnings relative to the participation thresholds.

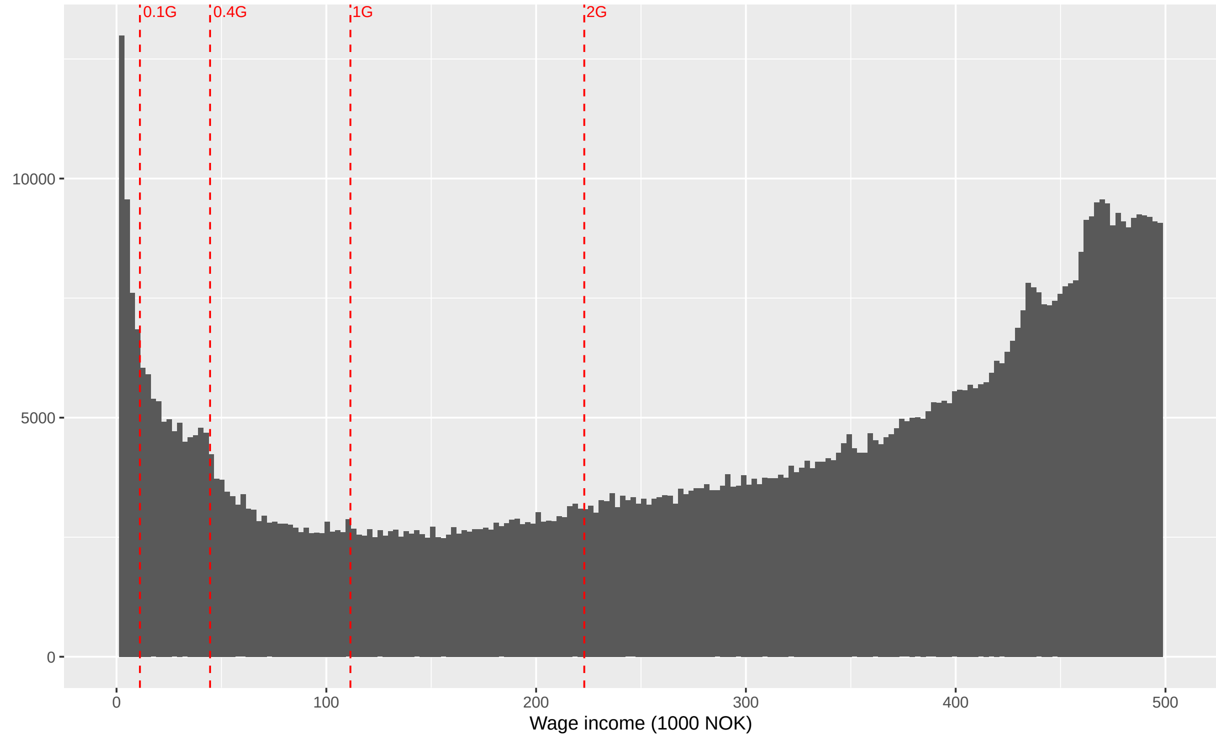
Table B2 shows that the extensive margin revenue effect increases as the threshold for participation increases. The number of people categorized as non-participants monotonically increases from around 400,000 in the 0.1 G alternative to around 600,000 in the 2 G alternative. Despite the 2 G alternative resulting in fewer individuals moving to participation (due to the reform), compared to as under 0.1 G, the revenue effect is larger as the wage income of those who transition into work is higher (on average), and this latter effect dominates with respect to the revenue effect.

Table B2 **Revenue effects of work deduction for different participation thresholds. Effects on revenue and switches for alternative counterfactual income in the state of non-participation**

Participation threshold	Extensive margin revenue effect	Switching individuals	Mean prediction wage (1,000 NOK)
0.1 G	463	4,134	211.2
	(447, 479)	(4,014, 4,254)	(50.4, 270.6)
0.4 G	792	6,005	266.3
	(769, 815)	(5847, 6164)	(111.1, 353.7)
1 G (baseline)	1,118	7,529	345.6
	(1,085, 1,142)	(7,364, 7,698)	(203.1, 438.7)
2 G	1,153	5,411	427.8
	(1,119, 1,185)	(5,254, 5,554)	(305.2, 503.8)

Notes: Top row shows median values (50%), bottom row shows confidence intervals (2.5%, 97.5%) for revenue and switching individuals, and quartile ranges (Q1, Q3) for mean predictions for wage earnings in 1,000 NOK. The 1 G alternative corresponds to the results presented in the main text.

Figure B1 **Thresholds for participation and the observed distribution of wage income**



Notes: This figure shows the distribution of wage income > 0, for the full population between the ages of 25 and 61. The vertical lines illustrate the different levels of wage income for categorizing a given person working/not working

B.3 Sensitivity to elasticity assumptions

As anticipated, the findings are sensitive to the assumed size of the response estimates. Tables B3 and B4 demonstrate this for the extensive margin effects of the bracket tax change and the work deduction, respectively. The extensive margin revenue effect increases approximately linearly with the response elasticity. For instance, when the elasticity rises from 0.1 to 0.5, the revenue effect becomes five times larger, consistent with a proportional (i.e., linear) relationship.

Table B3 **Revenue effects of five percentage-points increase in the two top brackets of the bracket tax. Behavioral effects at the extensive margin**

Extensive margin elasticity, λ	Revenue effect	Switching individuals
0.1	-328	406
	(-375, -288)	(362, 447)
0.2 (baseline)	-657	807
	(-715, -606)	(756, 868)
0.3	-986	1,215
	(-1,062, -915)	(1,140, 1,283)
0.5	-1,644	2,025
	(-1,754, -1,553)	(1,934, 2,116)

Notes: Top row shows median values (50%), bottom row shows confidence intervals (2.5%, 97.5%) for results across different elasticity values. All revenue values are negative, indicating revenue losses.

At the intensive margin, behavioral effects depend on the interaction between the compensated elasticity, ε^c, and the income effect, η'. Tables B5 and B6 report the resulting revenue effects under different elasticity assumptions. For the changes in the bracket tax, higher marginal tax rates combined with larger compensated elasticities lead to greater revenue losses, though the income effect partially offsets this. The results, however, are most sensitive to assumptions about the com-

Table B4 **Revenue effects of work deduction for different extensive margin elasticity estimates. Behavioral effects at the extensive margin**

Extensive margin elasticity, λ	Revenue effect	Switching individuals
0.1	557	3,767
	(536, 577)	(3,658, 3,883)
0.2 (baseline)	1,118	7,529
	(1,087, 1,142)	(7,364, 7,698)
0.3	1,670	11,301
	(1,636, 1,704)	(11,098, 11,504)
0.5	2,785	18,840
	(2,742, 2,830)	(18,582, 19,104)

Notes: Top row shows median values (50%), bottom row shows confidence intervals (2.5%, 97.5%) across different elasticity values.

pensated elasticity.

By contrast, for the work deduction, both the substitution effect and the income effect operate in the same direction, amplifying behavioral revenue losses as elasticities increase. Once again, the revenue estimates prove highly sensitive to the assumed behavioral responses.

Table B5 **Revenue effects of five percentage-points increase in the two top brackets of the bracket tax. Behavioral effects at the intensive margin for various intensive margin elasticity estimates**

	Income effect, η'			
	-0.01	-0.05	-0.2	-0.4
Compensated effect, ε^c		(baseline)		
0.05	-1,003	-856	-302	437
0.15 (baseline)	-3,079	-2,931	-2,377	-1,639
0.25	-5,148	-5,000	-4,446	-3,708
0.5	-10,291	-10,144	-9,591	-8,854

Table B6 **Revenue effects of work deduction. Behavioral effects at the intensive margin for various intensive margin elasticity estimates**

	Income effect, η'			
	-0.01	-0.05	-0.2	-0.4
Compensated effect, ε^c		(baseline)		
0.05	-572	-879	-2,026	-3,551
0.15 (baseline)	-1,555	-1,861	-3,010	-4,537
0.25	-2,527	-2,834	-3,983	-5,512
0.5	-4,922	-5,230	-6,381	-7,911