Statistisk sentralbyrå
Statistics Norway

Welfare effects of tax policy change when there are choice restrictions on labour supply

Zhiyang Jia and Thor O. Thoresen

TALL SOM FORTELLER

DISCUSSION PAPERS

959

Discussion Papers No. 959, August 2021
Statistics Norway, Research Department

Zhiyang Jia and Thor O. Thoresen

Welfare effects of tax policy change when there are choice restrictions on labour supply

Abstract:
Information about individual choices of heterogeneous agents. Results can for example be used to describe the distributional effects of tax policy change, such as the effects on changes in money metric utility – distributions of equivalent and compensating variation (*EV* or *CV*). This type of "revealed preference" methodology relies on using models with sufficient realism. In this paper we argue that the so-called "job choice model" represents a way forward in practical work, as it has a richer representation of choice constraints than conventional labour supply models. This model is also particularly suitable given an increased focus on distinguishing between preferences and constraints in applied welfare analysis. We demonstrate the empirical content of the framework by describing the effects of the Norwegian tax reform 2013–2019 on the distribution of compensating variation (*CV*).

Keywords: labour supply, money metric utility, distributional effects, tax reform

JEL classification: H31, I31, J22, C25

Acknowledgements: We have benefitted from discussions with and comments by John K. Dagsvik and Erling Holmøy. We are also grateful to comments by Andre Decoster, Norman Gemmell, Nicholas Hèrault, Guyonne Kalb, Christos Kotsogiannis, Alari Paulus and other participants of the "Modelling tax policy and compliance workshop", TARC, University of Exeter, 14–15 March 2019.

Address: Zhiyang Jia, Statistics Norway, Research Department. E-mail: jia@ssb.no

Thor O. Thoresen, Statistics Norway, Research Department and Oslo Fiscal Studies, Department of Economics, University of Oslo, E-mail: tot@ssb.no (Corresponding author)

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

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

Abstracts with downloadable Discussion Papers
in PDF are available on the Internet:
http://www.ssb.no/en/forskning/discussion-papers
http://ideas.repec.org/s/ssb/dispap.html

ISSN 1892-753X (electronic)

Sammendrag

Strukturelle arbeidstilbudsmodeller representerer viktige verktøy for å kunne beskrive individuelle beslutninger i arbeidsmarkedet. Slike modeller kan for eksempel anvendes til å beskrive fordelingseffekter av skatteendringer. Effekter av skatteendringer kan måles ved hjelp av pengemål på nytteendringene, som ekvivalent variasjon og kompensert variasjon. Sammenliknet med disponibel inntekt, som en ofte anvender i fordelingsanalyser av endringer i skattepolitikken, måler ekvivalent variasjon og kompensert variasjon effekter på individenes totale nytte. Det betyr at det tas hensyn til både endringer i forbruk og arbeidstilbud i målingen av effekter av skatteendringer. I dette arbeidet argumenteres det for at arebidstilbudsmodellen LOTTE-Arbeid (utviklet i SSB) er spesielt anvendbar til slike empiriske analyser fordi den, sammenliknet med standard arbeidstilbudsmodeller, gir en mer omfattende beskrivelse av individenes valg. I LOTTE-Arbeid modelleres individenes valg når det tas hensyn til at det er individuell variasjon i mulighetene på arbeidsmarkedet. Aktørene står overfor ulike muligheter på arbeidsmarkedet fordi det er forskjeller mellom individene i hvor mange jobber de tilbys. Siden det er økt fokus i litteraturen på velferdsmålinger der det skilles mellom valg som skyldes individenes preferanser og effekter av restriksjoner utenfor individenes kontroll, hevdes det at LOTTE-Arbeid er en modell som svarer på denne utfordringen. Det betyr at en i vurderingen av fordelingseffekter av skatteendringer kan skille mellom hva som primært tilhører individets eget ansvar, preferanser, og hva som skyldes ulikheter i rammebetingelser, som forskjeller i muligheter på arbeidsmarkedet. I dette arbeidet vises det hvordan LOTTE-Arbeid kan anvendes til å beskrive fordelingseffekter, målt ved kompensert variasjon, av de norske skatteendringene fra 2013 til 2019.

1 Introduction

The static structural labour supply model, with consumption and leisure as arguments in the utility function, represents a key source of information for predictions of the effects of tax policy change in the short-run, so-called “day after” effects. Models belonging to the Hausman type of labour supply models (Hausman, 1979), based on marginal calculus, or the discrete choice labour supply model (van Soest, 1995) can be used to predict or evaluate the effects of policy change on the distribution of well-being.

Although the effects of policy change on money metric utility occupy a key position in applied welfare analysis, policy change assessments often continue to be made in terms of changes in net or disposable income. There seems to be resistance among analysts to exploiting the full capacity of the labour supply model and using more comprehensive welfare metrics in studies of the effects of tax policy on the distribution of well-being, such as money metric utility. Exceptions include Aaberge, Dagsvik, and Strøm (1995), Kornstad and Thoresen (2006), Labeaga, Olivier, and Spadaro (2008), Dagsvik, Locatelli, and Strøm, (2009), Creedy, Hérault, and Kalb (2011), Creedy and Hérault (2012), Bargain, Decoster, Dolls, Neumann, Peichl, and Siegloch (2013) and Decoster and Haan (2015).

One possible reason behind the reluctance to use measures of money metric utility in applied work is that it brings to the surface the complications of comparing the welfare of heterogeneous agents. The static labour supply model enriches the analysis of welfare implications of policy change in that it facilitates for accounting for heterogeneity in labour supply choice, but heterogeneity is also a source to complication in the discussion of welfare implications. For example, how can we compare the welfare of a hard-working person with a low wage to a high-ability person with the same income level and much more leisure? Such questions have a long tradition in the theoretical social choice literature and have recently been addressed in studies of “equality of opportunity” (Roemer 1998) and in the “fair treatment” literature (Fleurbaey and Maniquet, 2011). Fleurbaey and Maniquet (2011) develop a normative theory of social justice based on a distinction between individual constraints (outside the individual responsibility), that may invoke calls for compensation, and preferences, such as strong preferences for leisure, which fall within the responsibility of the individual and therefore should not be compensated.

However, these advances of the theoretical literature on the understanding and conceptualisation of welfare do not seem to have been followed up by corresponding developments of modelling tools to be used in applied analysis. The main point of the present paper is that, given the preference-constraint dyad, a particular labour supply model, the so-called “job choice model” (Dagsvik, 1994; Dagsvik and Strøm, 2006; Dagsvik, Jia, Kornstad, and Thoresen, 2014; Dagsvik and Jia, 2016), is well-suited for practical work. The job choice model emphasises that individuals face restrictions when they choose jobs and hours of work and therefore represents a more realistic depiction of the actual choices of heterogeneous individuals than the conventional labour supply model. Moreover, using this model to simulate the effects of policy change facilitates for a deeper understanding of the interpretation of results along the preference-constraint dimension. For example, it facilitates controlling for differences in the number of options in the labour market across individuals (outside individual control). The effects of a tax policy change can then be discussed by neutralising individual characteristics that are "constraints" in the labour market, which bring us to estimates of changes in the distribution of money metric utility that can be characterised as primarily belonging to the preference part of individual choice. This configuration therefore produces distributional effects under a specification of fewer choice constraints, which can be compared to to other descriptions. Here, the job choice model is used to describe the distributional effects of the Norwegian tax reform 2013–2019 as measured by compensating variation (*CV*).

The job choice model is an extension of the conventional discrete labour supply model (van Soest, 1995). In contrast to the conventional discrete labour supply model, the job choice model is explicitly tailored to accommodate the fact that an agent in the labour market faces a latent set of job opportunities, where each "job" is characterised by fixed hours of work, wage rate and other latent non-pecuniary attributes (tasks to be performed in the job, location, social environment, etc.). The distribution of labour supply in the job choice model has a similar functional form to that of the conventional discrete choice labour supply model – the key difference is that it depends on a measure that captures the respective latent choice sets of job opportunities. It turns out that under the distributional assumptions of the random taste-shifters, the choice sets can be conveniently represented by the number of jobs for the hours of work in each case, denoted opportunity measure.

Accordingly, when the job choice model is used to simulate the effects of policy change, *EV* and *CV* become functions of the budget constraint, the wage rate, the opportunity measure and the individual-specific covariates. Consequently, the job choice model allows the analyst to assess the welfare effect of policy interventions where either a budget constraint (tax, wage rate or non-labour income) is changed or the opportunity measure is changed, or both. The computation of welfare measures in the job choice model is not straightforward because the utilities are stochastic, and therefore the corresponding welfare measures also become stochastic. Both simulation approaches (McFadden, 1999) as well as analytic approaches (Dagsvik and Karlström, 2005) have been developed to calculate the distribution of *EV* and *CV*, which we come back to in the following.[1]

[1] See also Herriges and Kling (1999) for a review of the state of the art as well as previous contributions.

The use of measures of money metric utility has been the topic of much criticism and controversy, see Slesnick (1998), Decancq, Fleurbaey, and Schokkaert (2015) and Fleurbaey and Blanchet (2013). There appears to be considerable scepticism regarding interpersonal comparison of well-being when individuals differ along several dimensions, such as having different preferences and facing dissimilar constraints. As a result, some studies have neglected preference heterogeneity at the normative stage of the analysis of money metric utilities, by letting all individuals have the same common preference ordering, that of a reference unit at reference prices; see King (1983).[2]

In the present analysis, however, we maintain preference heterogeneity at the welfare aggregation stage as well. Then a practical issue in this type of analysis concerns the choice of reference price. In fact, our contribution can be seen as alluding to the choice of reference prices for interpersonal comparison in analysis of welfare, as we propose to use the use the job choice model to specify a suitable welfare metric. In practical work, some authors continue using the individual-specific information in the welfare aggregation part of the analysis, whereas others argue that interpersonal comparison is achieved by using a common reference value (Preston and Walker, 1999; Decancq, Fleurbaey, and Schokkaert 2015; Decoster and Haan, 2015). The latter approach can be seen as establishing a measure of money metric utility based on finding an equivalent to the indirect utility of individual i, $v(y_i, p_i)$, where y_i is income and p_i is the individual-specific price. This means that we define a corresponding income equivalent, seen as $v(y_i^*, p^*)$, where p^* is the reference price. A major challenge is that this reference price may be arbitrarily chosen and that the ranking of social states need not be invariant to this choice (Slesnick, 1998). Preston and Walker (1999) and Decoster and Haan (2015) discuss possible choices of reference value and the welfare metrics implied.[3]

Fleurbaey and Maniquet (2011) argue that the choice of reference price does not need to be arbitrary; it can be guided by clear normative principles. The so-called fairness approach to social choice by Fleurbaey and Maniquet is helpful for ranking individuals when outcomes differ with respect to endowed circumstances and individual preferences. Our contribution to this discussion comes with a practical angle: we demonstrate how a reference situation can be further characterised by applying a more advanced labour supply model, controlling for individual differences with respect to options in the labour market. As such, the present analysis provides answers to requests for broader definition of reference sets (Fleurbaey and Blanchet, 2013; Fleurbaey, 2015).

The paper is organised as follows: The job choice model is presented in Section 2. Section 3 presents methods for obtaining measures of the welfare effect of tax policy change by applying a

---

[2] More fundamentally, Blackorby and Donaldson (1988) argue that social welfare functions that use money metric utility functions as arguments need not be quasi-concave, and therefore exhibit undesirable ethical properties.

[3] See also Carpantier and Sapata (2016).

discrete choice labour supply model, and we also describe how the job choice model, with its emphasis on choice set differences across individuals, can be applied. The empirical illustrations of the effects of policy change on measures of change in the money metric utility are presented in Section 4. Section 5 provides a conclusion for the paper.

2. The job choice model

It is a major theoretical and econometric challenge to develop models that are able to represent labour market behaviour in a reasonably realistic way, and at the same time are practical to apply in an empirical setting. The difficulty stems from the fact that budget constraints are nonlinear (piecewise linear) and may imply non-convex budget sets. This type of budget set follows from the tax systems in use in many countries, where deduction rules for different types of tax and withdrawal of social security payments at the lower income levels imply that marginal tax rates may not increase monotonically with income but may in fact decrease in specific income intervals.[4] Furthermore, it is complicated to apply the usual textbook approach based on marginal calculus unless simple labour supply functional forms are used (MaCurdy, Green, and Paarsch, 1990).

The use of discrete choice models has therefore become increasingly popular, as this approach drastically simplifies the implementation of complex nonlinear budget constraints, for instance for spouses in two-adult households (van Soest, 1995). Details of complex tax and transfer systems can easily be accommodated, and simulations of alternative policies can be carried out relatively straightforwardly. Examples of labour supply studies which are based on the discrete choice framework include Keane and Moffitt (1998), Blundell, Duncan, McCrae, and Meghir (2000), van Soest, Das, and Gong (2002), and Haan and Steiner (2005).

However, as we mentioned above, the conventional discrete labour supply model (van Soest, 1995) is silent about possible restrictions individuals are likely to face in the labour market. Indeed, in most countries observed hours of work figures show considerable peaks at full-time and possibly part-time hours of work, which indicates that most jobs only offer full-time or part-time hours of work schedules. This evidence has been the key motivation for developing the so-called job choice model, which is an extension of the conventional discrete labour supply model (Dagsvik, 1994; Dagsvik and Strøm, 2006; Dagsvik, Jia, Kornstad, and Thoresen, 2014; Dagsvik and Jia, 2016). This extension is based on the assumption that the fundamental choice variable is "job type", where the jobs are characterised by given hours of work, wage rate and other non-pecuniary job-specific attributes.

---

[4] Early contributions to the modelling of labour supply in a continuous choice context were made by Burtless and Hausman (1978) and Hausman (1979).

Furthermore, not all the jobs in the market are available to every individual, and most individuals are likely to face restrictions on their set of job possibilities and hours of work. It is this characteristic of the model which we argue in the present study can be used in applied welfare analysis.

We provide an outline of the job choice model. For simplicity we only consider households with a single decision maker. The extension to the case of married couples is briefly reviewed in the Appendix, see Section A.1. As the module for married couples is used to produce empirical illustrations in the present study, estimation results for couples are provided.

Households derive utility from household consumption, here set equal to household disposable income, leisure and non-pecuniary attributes of latent jobs. Let $z = 1, 2, ...,$ be an indexation of the jobs and let $z = 0$ represent the alternative of not working. The utility function is assumed to have the form

(2.1) $$U(C,h,z) = \log\mu + u(C,h) + \varepsilon(z,h) \quad \text{and} \quad U(C,0) = u(C,0) + \varepsilon(0)$$

where (C, h) denotes disposable income and annual hours of work, $u(C,h)$ is a suitable deterministic and continuous function that is common to all observationally identical individuals and $\varepsilon(z,h)$ is a stochastic error term (taste-shifter) that is supposed to capture unobserved heterogeneity of preferences relating to combinations of hours of work and job-specific attributes. Finally, $\log\mu$ is a constant that represents the "intrinsic" utility of being employed. This parameter permits a discontinuous jump of the utility when hours of work tend from the smallest positive value to zero. Here it is assumed that each job offers only a single given work schedule. Let $B(h)$ denote the set of jobs with hours of work h that are available to the individual. Further, let D be the set of feasible hours, including zero hours. The set $B(h)$, $h \in D$, is individual-specific and latent. The terms $\{\varepsilon(z,h)\}$ are assumed to be i.i.d. across individuals and across jobs for given h with the standard Gumbel cumulative distribution function $\exp(-e^{-x})$, for real x. The reason why the index z enters the utility function is that job-specific attributes other than wage and hours of work may affect the utility of the agents. For given hours of work h and wage rate w, household disposable income is given by $C = f(hw, y)$ where y is non-labour income and $f(\cdot)$ is the function that transforms gross income into disposable income.[5] Let $\theta g(h)$ denote the number of jobs in $B(h)$ where $g(h)$ is the fraction of jobs with hours of work h

[5] A more general formulation of the job choice model than the one assumed in this paper could allow for both intra- and inter-individual variations in wage rates. To this end, consider the general wage equation $\log W(z) = Rb + \lambda + \xi(z)$ where R is a vector of covariates, λ is an individual-specific random effect and $\xi(z)$ is an individual and job-specific random term that varies independently across jobs, with c.d.f. that is the same for all individuals. As discussed by Dagsvik and Jia (2016), one cannot achieve identification of this wage equation without strong functional form assumptions. In this paper, as well as in Dagsvik and Jia (2016) it is assumed that $\xi(z) = 0$ whereas Aaberge, Colombino, and Strøm (2004) assume that $\lambda = 0$. Dagsvik and Jia (2016) have tested the specification with $\xi(z) = 0$ against the specification with $\lambda = 0$ and they found that the former, with $\xi(z) = 0$, was rejected by their data whereas the latter was not.

that are available (to the individual) and θ is the total number of jobs that are available (to the individual). Here, both θ and $g(h)$ may depend on observed individual characteristics. Thus, $\theta g(h)$ is the opportunity measure referred to above,[6] which is the focus of attention in the present study. To extend the reference set (or equivalent income) we remove the individual-specific element of $\theta g(h)$, and obtain estimates of *CV* when this constraint does not differ across individuals.

From (2.1) it follows that the highest utility the agent can attain, given hours equal to *h*, is equal to

(2.2a) $$V(h,y) = \max_{z\in B(h)} U(f(hw,y),h,z) = \log\mu + u(f(hw,y),h) + \max_{z\in B(h)} \varepsilon(z,h)$$

and

(2.2b) $$V(0,y) = U(f(0,y),0) = u(f(0,y),0) + \varepsilon(0).$$

It follows immediately from (2.2a, b) and the Gumbel distributional assumption that the probability $\varphi(h)$ of choosing a job with hours of work equal to *h* becomes

(2.3a) $$\varphi(h) = P\Big(V(h,y) = \max_{x\in D} V(x,y)\Big) = \frac{\exp(u(f(wh,y),h))g(h)\theta\mu}{\exp(u(f(0,y),0)) + \mu\theta \sum_{x\in D\setminus\{0\}} \exp(u(f(wx,y),x)g(x)}$$

and

(2.3b) $$\varphi(0) = P\Big(V(0,y) = \max_{x\in D} V(x,y)\Big)\frac{\exp(u(f(0,y),0))}{\exp(u(f(0,y),0)) + \mu\theta \sum_{x\in D\setminus\{0\}} \exp(u(f(wx,y),x)g(x)}.$$

The distribution assumptions about the random error terms $\{\varepsilon(z,h)\}$ can be justified by the notion of probabilistic rationality (Luce, 1977) expressed by the Choice Axiom of Luce. It has the fortunate implication that the choice sets of job opportunities can be represented by the respective number of jobs in the respective choice sets $\{B(h)\}$. Given the information available in most datasets, the term μ cannot be identified separately from θ and it can therefore be normalized to 1. This implies that the estimate of θ can no longer be interpreted as the number of jobs available to the worker, but it can still be interpreted as a measure of job availability since μ is a constant preference factor.

[6] We do not discuss the formation of the choice sets of available jobs. Dagsvik (2000) and Dagsvik and Jia (2018) consider the job choice model in a two-sided matching setting in which workers are searching for a suitable job and firms are looking for suitable workers. Under quite general conditions on the distribution of stochastic preferences, it turns out that the resulting choice model of the workers in large labour markets has the same structure as the job choice model above. Dagsvik and Jia (2018) demonstrate that the sizes of the endogenous and stochastic choice sets of jobs converge towards deterministic terms when the labour market becomes large.

3. Welfare measurement

3.1. Stochastic measures of money metric utility

In order to conduct welfare analyses we need to obtain welfare measures (*EV* or *CV*) for the job choice model. To this end, we first need to derive expressions for the corresponding indirect utility function. Due to the distributional assumptions of the random error terms, we show in the Appendix (Section A.2), that the conditional indirect utility given *h* (and given the opportunity sets) is equivalent to (has the same distribution as)

(3.1a) $$V(h,y) = u(f(hw,y),h) + \log(\theta g(h)) + \eta(h)$$

for $h > 0$, where $\eta(h)$ is the error term, and

(3.1b) $$V(0,y) = u(f(0,y),0) + \varepsilon(0) = u(f(0,y),0) + \eta(0)$$

for $h = 0$, where $\varepsilon(0), \eta(1), \eta(2), \ldots,$ are i.i.d. with standard Gumbel c.d.f. It is indeed a convenient property of the Gumbel c.d.f. that the distribution of $\{\eta(h)\}$ turns out to be the same as the distribution of $\{\varepsilon(z,h)\}$. Thus, we see from (3.1a) that the indirect utility function differs from the corresponding indirect utility in the standard discrete choice model by the logarithm of the opportunity measure. We realise that the calculation of *EV* and *CV* in the job choice model can be done in the same way as in the conventional discrete labour supply model, but with the addition of a term representing job options, $\log(\theta g(h))$.

In the empirical model (see Section A.1 in the Appendix) the deterministic part of the utility function and the opportunity measure depend on individual characteristics. Specifically, θ may depend on schooling and gender whereas $g(h)$, which is thought to represent overall institutional constraints, is determined by the nature of the tasks to be performed and by negotiations between labour unions. Since the job model accounts for preference heterogeneity as well as heterogeneity in job opportunities it opens for the possibility of computing welfare measures that imply compensation for differences in job opportunities.

Consider now a policy change in the budget set and the opportunity measure where the ex-ante function f is changed to the ex post function $\tilde{f}$ and the opportunity measure and wage rate are changed from $(\theta g(h), w)$ to $(\tilde{\theta}\tilde{g}(h), \tilde{w})$. Define

$$v(C,h,\theta g(h)) = u(C,h) + \log(\theta g(h))$$

when *h* is positive and

$$v(C,0,\theta g(0)) = u(C,0)$$

when *h* = 0. Using this notation and (3.1a) and (3.1b), the *CV* measure for individual *i* is defined by

(3.2) $$\max_{h\in D}(v_i(f(w_i h, y_i), h, \theta_i g_i(h)) + \eta_i(h)) = \max_{h\in D}(v_i(\tilde{f}(\tilde{w}_i h, \tilde{y}_i) + CV_i, h, \tilde{\theta}_i \tilde{g}_i(h)) + \eta_i(h))$$

where the indexation of the individual refers to the individual-specific covariates that are suppressed in the notation above. This definition of *CV* is analogous to the standard definition; see McFadden (1999). It follows that *CV* is stochastic because it depends on all the stochastic terms $\{\eta_i(h)\}$ and that it incorporates the opportunity measure.[7] Section A.4 in the Appendix shows how measures of *EV* and *CV* can be obtained under this framework.[8]

3.2. Welfare comparisons with heterogeneous tastes

Fleurbay and Maniquet (2011) demonstrate how to derive individual welfare metrics in terms of money metric utility in a normative framework.[9] They argue that reference sets should be selected on the basis of clear normative principles, or at least the normative choices should be made explicit (Decancq, Fleurbaey, and Schokkaert, 2015).[10] Our take on this challenge is to direct attention to model realism for obtaining welfare metrics; as information is obtained from revealed preferences, a rich and realistic model is beneficial. We argue that the job choice model represents a way forward in this respect.

The money metric utility or equivalent income is defined as the amount of money hypothetically needed under reference prices that would give everyone their current level of satisfaction (Slesnick, 1998; Fleurbaey, 2015). In the following, we illustrate the contribution of the present paper in such a framework. Let $v_i(y_i, p_i, q_i)$ denote the indirect utility obtained by individual i with income y_i, prices p_i and other conditions q_i. As in Fleurbaey (2015), the comparable money metric utility, y_i^* must satisfy $v_i(y_i, p_i, q_i) = v_i(y_i^*, p^*, q^*)$, where p^* is a reference price vector and q^* represents other dimensions that one may also take into account. In other words, to obtain y_i^*, we need to assign a common wage to each individual, thereby setting a reference price vector (symbolised by p^*), and more importantly, assigning a common q^*. Given that our labour supply model holds properties that facilitate controlling for important dimensions in the compensation-responsibility dyad in the discussion on fairness of taxation, we maintain that applying the job choice model in practical work is an example of bringing empirical content to analysis along these lines.

---

[7] There are no closed-form formulae for *EV* and *CV*. However, Dagsvik and Karlström (2005) give formulae for the distribution and moments of *CV*. Section A.3 in the Appendix summarises this approach.

[8] Another empirical concern is how the change in the money metric measure, in this case *CV*, should be aggregated in a population. We return to this in the discussion of empirical results in Section 4.

[9] However, the approach of Fleurbaey and Maniquet (2011) is characterized as non-welfarist, as it relies on fairness principles.

[10] See also Piacquadio (2017) and Bosmans, Decancq and Ooghe (2018) on the aggregation of money metric utility.

More precisely, we redefine Equation (3.2), to compute *CV* due to a change in the tax system,[11] by controlling for individual heterogeneity in wage rates and opportunities by introducing a reference wage rate w^0 (say) and a reference opportunity measure $\theta^0 g^0(h)$ (say). A common reference wage means that inequalities arising from inequalities in ability is removed, which can be argued belong to circumstances outside the individual's control.[12] The control for differences in opportunities in the labour market, represented by $\theta^0 g^0(h)$, adds to controlling for conditions outside the control of individuals. The corresponding *CV* measure now becomes

$$(3.3) \qquad \max_{h\in D}(v_i(f(w^0h,y_i),h,\theta^0 g^0(h))+\eta_i(h)) = \max_{h\in D}(v_i(\tilde{f}(w^0h,y_i)+CV_i,h,\theta^0 g^0(h))+\eta_i(h)).$$

It follows that the resulting measure of change in money metric utility therefore gets closer to representing preferences, and the distribution of *CV* derived from Equation (3.3) can be interpreted accordingly.

4. Welfare effects of the Norwegian tax reform 2013–2019

4.1 The Norwegian tax reform 2013–2019

In this section we describe the distributional effects of the Norwegian tax reform 2013–2019 by discussing measures of the distribution of *CV* distribution. The Norwegian tax reform 2013–2019 followed two major reforms of the tax system in 1992 and 2006. The former introduced the dual income tax system in Norway, while the latter maintained the system with important modifications; see Sørensen (2005). Notably, the post-1992 tax system has adhered to the principle that the tax rate on personal capital income should equal the corporate tax rate. Throughout the period from 1992 to 2014, the tax rate applied to both tax bases was 28 percent. The tax schedule for labour earnings has been designed as consisting of the same basic tax rate as that on capital income, supplemented by a surtax schedule with several brackets and a 7.8 percent national insurance contribution.

The motivation for the 2013–2019 reform is different from the considerations underlying the 1992 and 2006 reforms. It is the level of the basic tax rate applicable to capital and labour income and business profits that creates pressure now. A tax on corporate profits of 28 percent was low at the time of its introduction in 1992 (Ministry of Finance, 2015), but in subsequent decades Norway was left behind in the international tax competition (Devereux, Griffith, and Klemm, 2002), ending up with statutory tax rates well above the average in the OECD and in EU-28 (Ministry of Finance, 2015). To

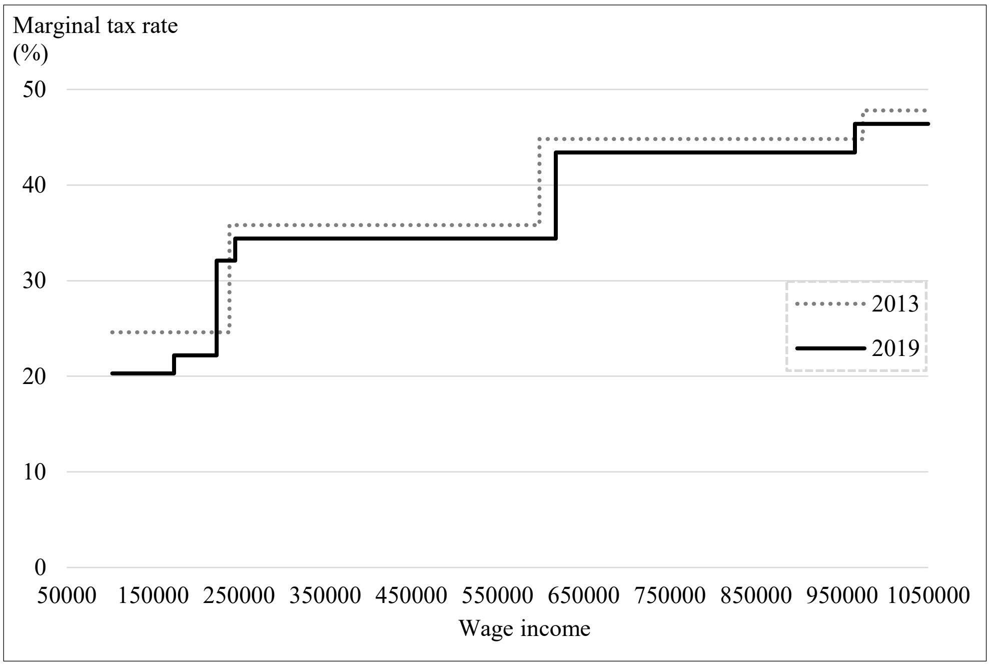
---

[11] We focus on estimates of *CV* in the following; similar estimates of *EV* are straightforwardly obtained by the same method.

[12] Of course, it can be argued that the wage is determined by factors that can be viewed as being under the individual's control, such as effort. This raises fundamental questions concerning the distinction between "efforts" and "endowments", which is beyond the scope of the present analysis.

improve the international tax position of Norway, the main element of the 2013–2019 reform was a gradual lowering of the corporate tax rate. Due to the link to the basic tax rate on general income of persons, following from the design of the Norwegian version of the dual income tax, the immediate consequence was a cut in the flat-rate part of the tax on labour earnings as well as the tax rate on personal capital income, down to 22 percent in 2019. Together these tax rate cuts obviously implied a significant loss of tax revenue, and a major challenge was to offset at least a substantial part of the foregone revenue. The main move has been to introduce more steps in the schedule for the progressive part of the tax on labour income, previously referred to as the surtax schedule. To distinguish the new schedule from the old one, the term "step-tax" was adopted to reflect the larger number of steps in the new step-wise linear income tax.

Figure 1. Marginal tax rates, 2013 and 2019



In Figure 1 the new schedule (as of 2019) is compared to the schedule of 2013. The figure clearly illustrates that there are two additional steps in 2019, compared to 2013, and that the 2019 rates in general are lower than those of 2013. The first step generates revenue at lower levels by introducing the first step, an addition of 1.9 percentage points, whereas the next step (2.3 percentage points) kicks in when the basic allowance is exhausted, which brings the marginal tax rate in the interval from approximately 230,000 Norwegian kroner (NOK)[13] to around NOK 600,000 very close to the pre-

[13] This corresponds to approximately 23,500 euros and 26,000 U.S. dollars according to average exchange rates for 2019: 1€=NOK9.85 and 1$=NOK8.80.

reform schedule, see Figure 1. The two last steps basically replicate the two-tier surtax schedule of the pre-reform system, with somewhat lower rates.

4.2 Distributions of CV

In the following we discuss the welfare effects of the tax reform 2013–2019 by obtaining distributions of *CV*, applying the measure of *CV* introduced in Section 3 and using the simulation approach of McFadden (1999); for the latter, see Appendix 4. The empirical results are derived by means of a discrete choice labour supply simulation model developed for Norwegian policy-makers under the name LOTTE-Arbeid (Aasness, Dagsvik, and Thoresen, 2007), which is based on the job choice framework (see Section 2).[14] We focus on the labour supply of married/cohabiting couples. The couples are unitary, which means that they make harmonised decisions with respect to the labour supply of each spouse, given a common budget constraint; see more details on practical implementation in Dagsvik, Jia, Kornstad, and Thoresen (2014), Thoresen and Vattø (2015) and Dagsvik and Jia (2016). We let the female of the couple choose between 8 different hours of work, whereas males have 7 discrete choices, which gives the couple a total of 56 possibilities. Given that the tax reductions of the reform make most taxpayers better off after the reform, measures of *CV* are negative. It follows from this that the taxpayers benefitting most from the reform have large negative money metric compensations.

In the discussion of results, we refer to descriptions of three different simulation alternatives. First, we obtain measures of welfare change when using individual-specific information to establish the reference set, similar to what has been derived previously by means of discrete choice models; see, for example, Aaberge, Dagsvik, and Strøm (1995), Labeaga, Olivier, and Spadaro (2008), and Dagsvik, Locatelli, and Strøm (2009). We characterize this as the benchmark. Second, we obtain estimates for distributions of *CV* for two definitions of the reference set (see Section 3.2), which are compared to the benchmark distribution: one alternative where we assign a fixed identical wage to males and females, and a second where we let labour market opportunities across families be equalised as well. As discussed in Section 2, it is possible to obtain estimates for the latter alternative because of an idiosyncratic characteristic of the job choice model, which provides individual measures of labour market opportunities as its output. To assess the effects of the different stages of the analysis, we address differences in the distribution of *CV* for the different simulations. Ultimately, the ambition is to end up with a measure of *CV* which results largely from individual preferences, here operationalised by assigning an identical wage to the individuals and equalising the number of options in the labour

[14] This partial perspective means (of course) that general equilibrium effects are left out of the analysis.

market. Distributional effects according to these variations in specifications are informative about distributional effects from a preference-constraint perspective, answering to the challenge of, for example, Fleurbaey and Maniquet (2011) to distinguish between preferences and matters outside individual control.

Table 1 describes summary statistics for the three alternative simulations: first, for the benchmark simulation based on employing individual-specific information for both wages and options in the labour market (referred to as "No reference set restriction" in Table 1); second, for a simulation in which wages are fixed by gender; and third, for a simulation in which both wages (for males and females) and options in the labour market ($\theta g(h)$) are fixed across individuals. Recall that *CV* measures the income needed to restore an individual's original level of utility following the policy change. Accordingly, as most people benefit from the reform, the compensations become negative. As we use McFadden's simulation approach, based on draws for the random part of the utility function (see Appendix A.4), we obtain expected (or mean) values of *CV* for each household, which are input for the calculation of population means in Table 1.[15]

The distributional effects of the reform in terms of *CV* are described by four different graphs for each of the three different simulations. Figure 2 (for the benchmark case) serves as an example for the informational content. In Figure 2, the frequency distribution is described first, in the upper left-hand diagram. We find measures of *CV* from 0 to around NOK -75,000, with the majority in the range NOK -5,000 to NOK -20,000. Next, distributions of *CV* are described against non-labour income, disposable income, age and years of schooling.[16] Of course, another possibility is to describe *CV* distributions when the individuals are ranked according to equivalent income (or money metric utility). We do not show that type of result here, restricting ourselves to using money metric utility as a measure of the effect of tax reform and using other measures of well-being, such as disposable income, for ranking individuals.

Table 1. Summary statistics for simulation results

Simulation	$\mathrm{E}(CV)$	Interquartile range, NOK
No reference set restriction	-66,027	11,738
Fixed wage	-62,927	1,598
Fixed wage, fixed $\theta g(h)$	-63,921	1,518

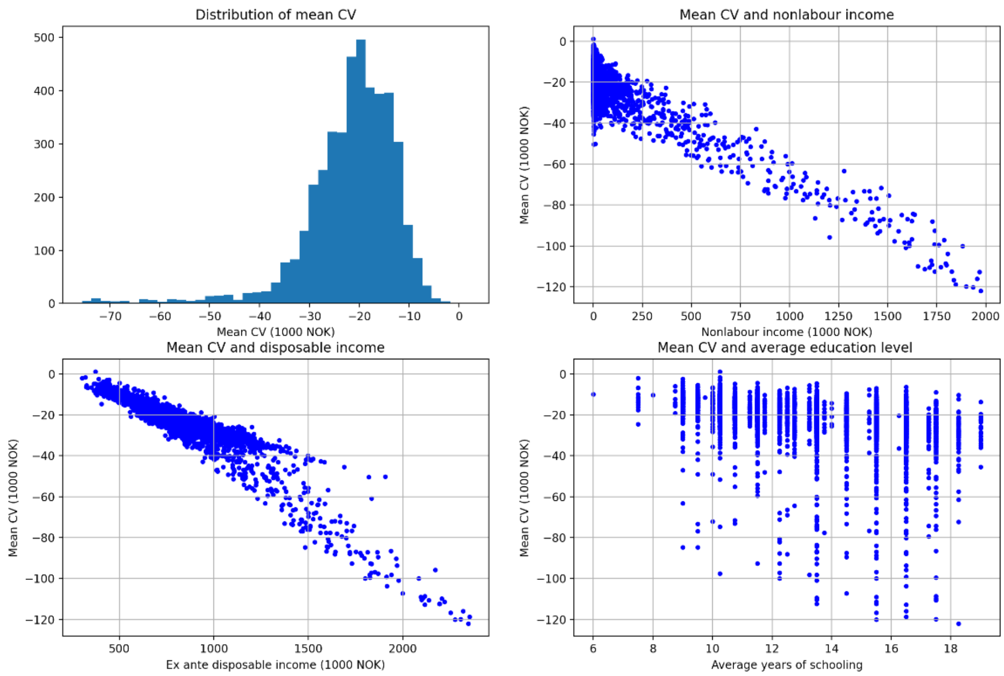
Notes. Values in Norwegian kroner

[15] The average values reported in Table 1 are substantially above the peaks of the distributions (which we return to soon), due to some large (negative) outlier observations.

[16] For couples we use the average across partners.

The results shown in Figure 2 reflect the fact that the tax reform involved substantial reductions in capital income taxation, which is included in non-labour income. The tax reform benefits those with high non-labour income to a greater extent; see top right-hand side diagram of Figure 2, i.e., those with high non-labour income are willing to pay most for the reform, and thus have the largest negative compensation. This result is also applicable to disposable income (middle left-hand side diagram); the largest negative compensation is seen for the high end of the distribution of disposable income.[17] The relationship to age is less clear, whereas years of schooling increase with (negative) *CV*.

Figure 2. Effects on *CV* of the Norwegian tax reform, 2013–2019. Observed individual wage and individual job opportunities

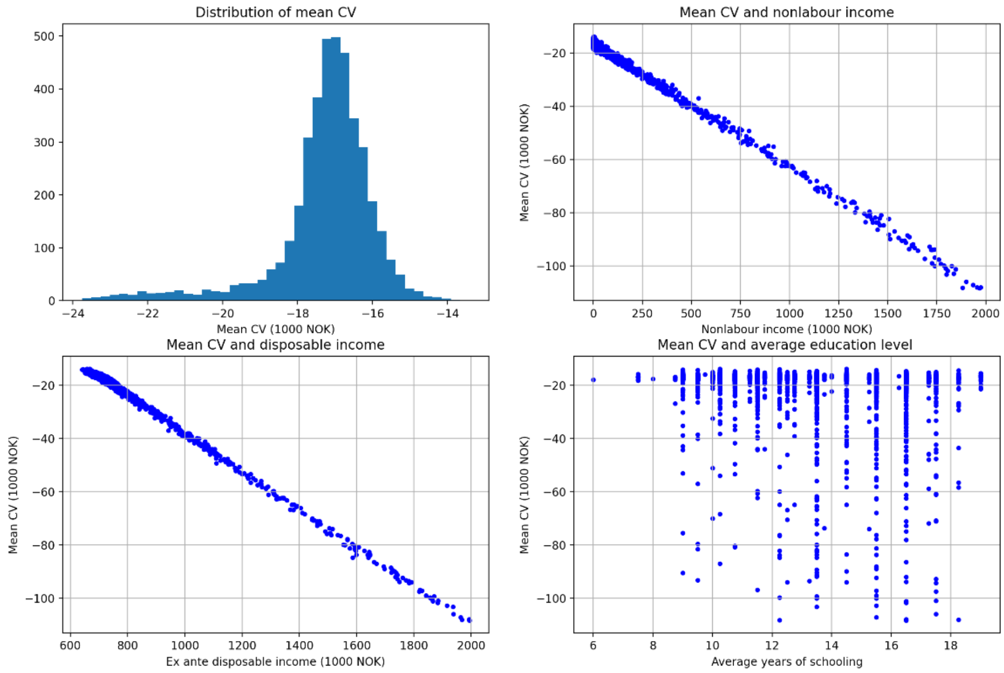


Notes. The figure contains four graphs, each showing a *CV* distribution. The density distribution of *CV* is shown top left. The graphs at top right and bottom left show the distribution of *CV* against non-labour income and disposable income, respectively. The bottom right-hand graph shows the relationship between *CV* and education. As estimates of *CV* are based on random draws, we obtain mean values for each household.

[17] These results are in line with Christiansen, Jia and Thoresen (2021), who find that the reform can be justified (it gives an overall welfare gain) only if moderate inequality aversion prevails. Results are derived under the assumption of a trade-off between allocative efficiency gains and redistributive losses.

Next, Figure 3 shows distributional effects when the first step of establishing a reference set for interpersonal comparison is taken: the wage rates for females and males are fixed as NOK 275 for males and NOK 255 for females. The figure reveals that the main implication of assigning the same wage to everyone is that the variability in *CV* for the same level of non-labour income and disposable income becomes much smaller. This is expected, given that a substantial source of variation, differing wages, is removed. It follows that the main pattern of Figure 2, in which *CV* increases with respect to both non-labour income and disposable income, becomes even more pronounced in Figure 3.

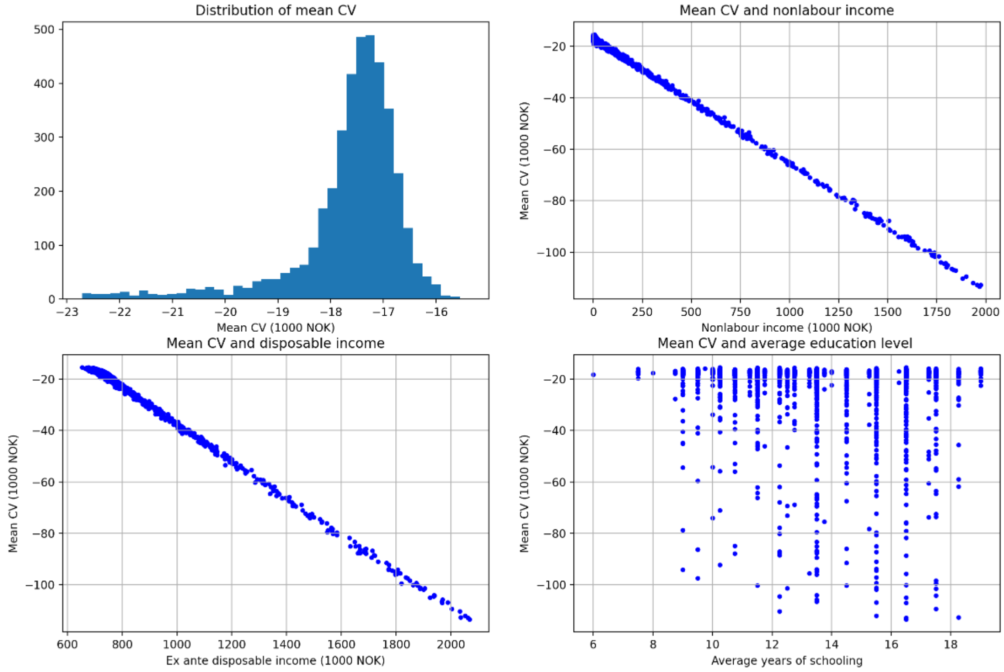
Figure 3. Effects on *CV* of the Norwegian tax reform, 2013–2019. Identical gender-specific wage and individual job opportunities



Notes. The figure includes four graphs, each showing a distribution of *CV*. The density distribution of *CV* is shown top left. Top right and bottom left-show the distribution of *CV* against non-labour income and disposable income, respectively. The graph at bottom right shows the relationship between *CV* and education. As estimates of *CV* are based on random draws, we obtain mean values for each household.

Finally, Figure 4 shows the distributional effects of adding the second part of our reference set formulation: the cancellation of individuals with different options in the labour market; see Equation (3.3). In practice this is done by letting all individuals face the same number of options in the labour market corresponding to having 13 years of education (given that options in the labour market increase in years of education). As wage rates are fixed and gender-specific and job options do not vary across individuals, the resulting *CV* distribution depicts a distribution of the welfare effects of the reform that primarily reflect preferences, as we have operationalised it here. This simulation implies that two characteristics assigned to "outside individual control", i.e. the wage and the number of job options in the labour market, are controlled for.

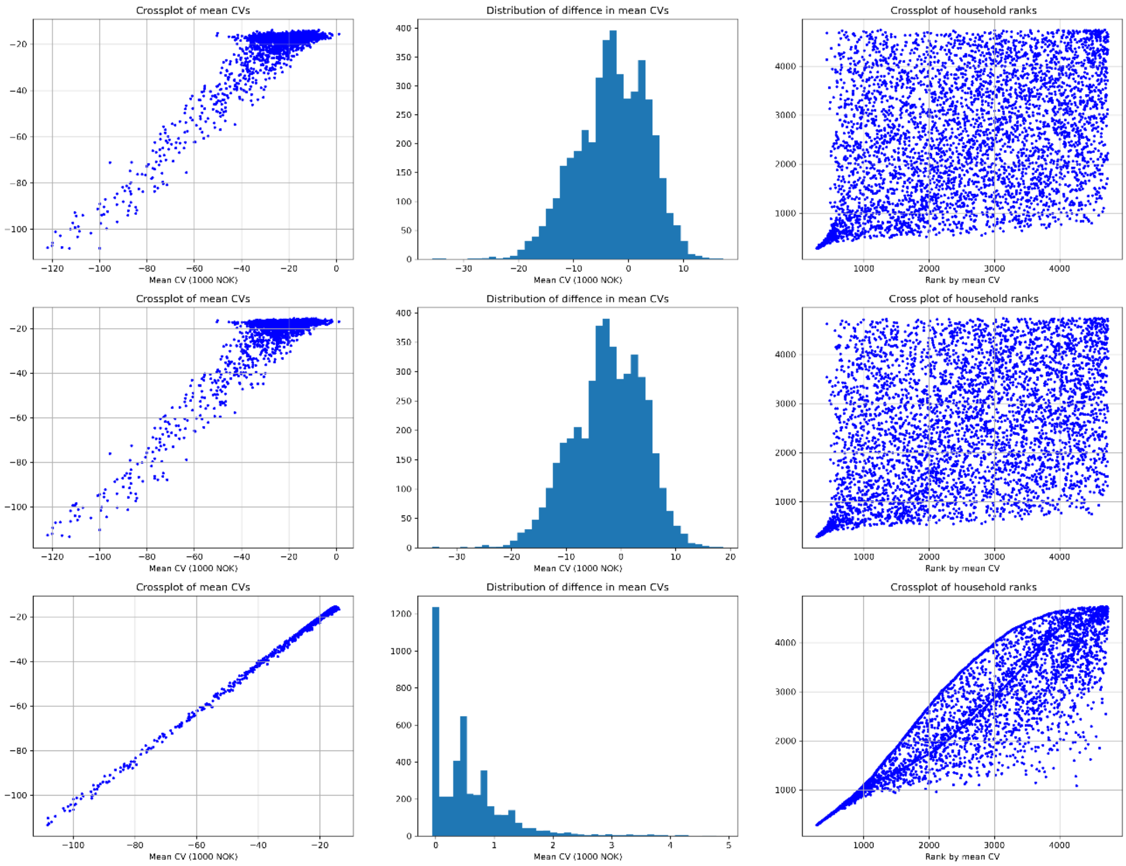
Figure 4. Effects on *CV* of the Norwegian tax reform, 2013–2019. Identical gender-specific wage and identical job opportunities



Notes. The figure includes four graphs, each showing distributions of *CV*. The density distribution of *CV* is shown top left. Top right and bottom left graphs show the distribution of *CV* against non-labour income and disposable income, respectively. The graph at bottom right shows the relationship between *CV* and education. As estimates of *CV* are based on random draws, we obtain mean values for each household.

The results of Figure 4 do not seem to deviate much from what we see in Figure 3, but there are important differences between the two simulation results. This is demonstrated by Figure 5, which shows differences between pairs of simulations; cross-plots of *CV*, distribution of difference in *CV* according to the simulations and cross-plots of household rankings. Thus, each row of Figure 5 describes pair-wise differences in simulation results for the following three simulations: individual-specific wage rates and individual-specific labour market opportunities (reported in Figure 2); gender-specific wage rates and individual-specific labour market opportunities (Figure 3); and gender-specific wage rates and equal number of job offers in the labour market (Figure 4). This means that in the upper row of Table 5 we compare results reported in Figure 2 and Figure 3, the middle row shows differences between Figure 2 and Figure 4, and the bottom row shows differences between Figure 3 and Figure 4.

Figure 5. Comparison of three simulations



Notes. The three rows in the figure describe comparisons between three simulations. From left to right: crossplot of *CV*, distribution of difference in *CV*, and crossplot of household rankings. In the upper row of the table, the benchmark alternative is compared to the common wage (gender-specific) alternative, the middle row shows a comparison of the benchmark alternative and the alternative with both common wage and common labour market options, and the bottom row shows a comparison of the common wage alternative and the alternative with both common wage and common labour market.

The middle graph of the bottom row of Figure 5 shows that when wages have already been equalised, neutralising the number of job options has an additional effect. Although the choices of many individuals do not change, the difference in *CV* is between NOK 1 and NOK 2,000 for a large number of people. This also means that the ranking of individuals in terms of *CV* depends on the definition of the reference set, which is shown by the bottom right graph in Figure 5.

The goal of this exercise is to obtain reform results based on a configuration which primarily generates preference-driven measures of *CV*. The results presented in Figure 4 represent this, showing the effects of the 2013–2019 Norwegian tax reform in terms of distributions of a measure of *CV* resulting from removing characteristics defined as outside individual control. These results therefore exemplify how to proceed in order to develop a practical framework for discussing the effects of tax reforms in a "fair tax treatment" perspective. As such, the present analysis provides a way forward in empirical work, given the preference-constraint dichotomy of the new fairness literature, for example as developed by Fleurbaey and Maniquet (2011).

5. Conclusion

The change in money metric utility is a widely accepted means of describing the effects of tax reforms on individual utility and other changes in the decision-making environment of economic agents. Inspired by recent discussions in the social justice literature, we suggest that a particular discrete choice model, under the name "job choice model", can be used in practical work. Despite an increasing focus on characterising welfare in terms of money metric utility, which paves the way for differentiating between individual choice circumstances and preferences, we see less attention being paid to empirical work along these lines. For example, the advances of Fleurbaey and Maniquet (2011) do not seem to have been followed up by corresponding developments on the empirical side.

The present paper argues that the job choice model is tailored to address the distinction between responsibilities and constraints, demonstrating how such a tool can be used to provide information about the distribution of welfare effects of a policy change, such as a tax reform. As for the standard discrete choice model, the job choice model facilitates for accounting for heterogeneity in labour supply choice. By neutralising differences in individual wages and differences between individuals with respect to labour market options, it is argued that we obtain estimates of *CV* which can be characterised as primarily belonging to the preference part of individual choice. Thus, the present paper demonstrates how one may obtain empirical counterparts to recent theoretical advances in welfare analysis, such as those of Fleurbaey and Maniquet, where the distinction between preferences and constraints is pivotal.

When this is applied to the Norwegian tax reform of 2013–2019, we get negative *CV* values, as most individuals benefit from the reform. Negative *CV* values increase for both disposable income and non-labour income, and this pattern does not change when measures of *CV* reflecting preferences alone are obtained, in the manner defined and operationalised here. The distribution of a "preference-assigned" *CV* shows more precise relationships between *CV* and disposable income/non-labour income compared, as seen, to the distribution of *CV* when both wages and job options are individual-specific.

References

Aaberge, R., U. Colombino, and S. Strøm, (2004). Do more equal slices shrink the cake? An empirical investigation of tax-transfer reform proposals in Italy. *Journal of Population Economics*, 17(4), 767–785.

Aaberge, R., J. K. Dagsvik, and S. Strøm 1995. Labor supply responses and welfare effects of tax reforms. *Scandinavian Journal of Economics*, 97(4), 635–659.

Aasness, J., J. K. Dagsvik, and T. O. Thoresen (2007). "The Norwegian tax-benefit model system LOTTE". In A. Gupta and A. Harding (eds.), *Modelling Our Future: Population Ageing, Health and Aged Care*, *International Symposia in Economic Theory and Econometrics*. Amsterdam: Elsevier, pp. 513–518.

Bargain, O., A. Decoster, M. Dolls, D. Neumann, A. Peichl, and S. Siegloch (2013). Welfare, heterogeneity and responsibility: Evidence from Europe. *Social Choice and Welfare*, 41(4), 789–817.

Blackorby, C. and D. Donaldson (1988). Money metric utility: A harmless normalization? *Journal of Economic Theory*, 46(1), 120–129.

Blundell, R. W., A. Duncan, J. McCrae, and C. Meghir (2000). The labour market impact of the working families' tax credit, *Fiscal Studies*, 21(1), 75–103.

Bosmans K., K. Decancq, and E. Ooghe (2018). Who's afraid of aggregating money metrics? *Theoretical Economics* 13, 467–484.

Burtless, G. and J. Hausman (1978). The effect of taxation on labor supply: Evaluating the Gary Income Maintenance Experiment. *Journal of Political Economy*, 86(6), 1103–1130.

Carpantier, J.-F. and C. Sapata (2016). Empirical welfare analysis: when preferences matter. *Social Choice and Welfare*, 46(3), 521–542.

Christiansen, V., Z. Jia, and T. O. Thoresen (2021). Assessing income tax perturbations. *International Tax and Public Finance*, forthcoming.

Creedy, J. and N. Hérault (2012). Welfare-improving income tax reforms: A microsimulation analysis. *Oxford Economic Papers*, 64(1), 128–150.

Creedy, J., N. Hérault, and G. Kalb (2011). Measuring welfare changes in behavioural microsimulation modelling: Accounting for the random utility component. *Journal of Applied Economics*, 14(1), 5–34.

Dagsvik, J. K. (1994): Discrete and continuous choice, max-stable processes, and independence from irrelevant attributes. *Econometrica*, 62(5), 1179–1205.

Dagsvik, J. K. and A. Karlström (2005). Compensating variation and Hicksian choice probabilities in random utility models that are nonlinear in income. *Review of Economic Studies*, 72(1), 57–76.

Dagsvik, J. K., M. Locatelli, and S. Strøm (2009). Tax reform, sector-specific labor supply and welfare effects. *Scandinavian Journal of Economics*, 111(2), 299–321.

Dagsvik, J. K. and Z. Jia (2016). Labor supply as a choice among latent jobs: Unobserved heterogeneity and identification. *Journal of Applied Econometrics*, 31(3), 487–506.

Dagsvik, J. K. and Z. Jia (2018). Aggregate behavior in matching markets with flexible contracts and non-transferable representations of preferences. Discussions Papers No. 875, Statistics Norway.

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(1), 134–151.

Dagsvik, J. K. and S. Strøm (2006). Sectoral labor supply, choice restrictions and functional form, *Journal of Applied Econometrics*, 21(6), 803–826.

Decoster, A. M. J. and P. Haan (2015). Empirical welfare analysis with preference heterogeneity. *International Tax and Public Finance*, 22(2), 224–251.

Decancq K., M. Fleurbaey, and E. Schokkaert (2015). "Inequality, income, and well-being." In A. B. Atkinson and F. Bourguignon (eds.), *Handbook on Income Distribution*, vol 2A. Amsterdam: Elsevier, pp. 67–140.

Devereux, M. P., R. Griffith, and A. Klemm (2002). Corporate income tax reforms and international tax competition. *Economic Policy*, 35, 451–495.

Fleurbaey, M. (2015). Beyond income and wealth. *Review of Income and Wealth*, 61(2), 199–219.

Fleurbaey, M. and D. Blanchet (2013). *Beyond GDP. Measuring Welfare and Assessing Sustainability*. Oxford: Oxford University Press.

Fleurbaey, M. and F. Maniquet (2011). *A Theory of Fairness and Social Welfare*. Cambridge University Press: Econometric Society Monographs.

Haan, P. and V. Steiner (2005). Distributional effects of the German tax reform 2000 – a behavioral microsimulation analysis. *Journal of Applied Social Science Studies*, 125, 39–49.

Herriges, J. A. and C. L. Kling (1999). Nonlinear Income Effects in Random Utility Models. *The Review of Economics and Statistics*, 81(1), 62–72.

Hausman, J. A. (1979). The econometrics of labor supply on convex budget sets. *Economics Letters*, 3(2), 171–174.

Keane, M. and R. Moffitt (1998). A structural model of multiple welfare program participation and labor supply. *International Economic Review*, 39(3), 553–589.

King, M. A. (1983). Welfare analysis of tax reforms using household data. *Journal of Public Economics*, 21(2), 183–214.

Kornstad, T. and T. O. Thoresen (2006). Effects of family policy reforms in Norway: Results from a joint labour supply and childcare choice microsimulation analysis. *Fiscal Studies*, 27(3), 339–371.

Labeaga J. M., X. Oliver, and A. Spadaro (2008). Discrete choice models of labour supply, behavioural microsimulation and the Spanish tax reforms. *Journal of Economic Inequality*, 6(3), 247–273.

Luce, R. D. (1977). The choice axiom after 20 years. *Journal of Mathematical Psychology*, 15, 215–233.

MaCurdy, T., D. Green, and H. Paarsch (1990). Assessing empirical approaches for analyzing taxes and labor supply. *Journal of Human Resources*, 25, 414–490.

McFadden, D. (1999). "Computing willingness-to-pay in random utility models". In J. Moore, R. Riezman, and J. Weeks (eds.), *Trade, Theory and Econometrics: Essays in Honour of John S. Chipman*, Routledge Studies in the Modern World, Vol. 15. London: Routledge, pp. 253–274.

Ministry of Finance (2015). Meld.St. 4. Bedre skatt. En skattereform for omstilling og vekst. (In Norwegian)

Piacquadio, P. G. (2017). A fairness justification of utilitarianism. *Econometrica*, 85(4), 1261–1276.

Preston, I. and I. Walker (1999). Welfare measurement in labour supply models with nonlinear budget constraints. *Journal of Population Economics*, 12(3), 343–361.

Roemer, J. E. (1998). *Equality of Opportunity*. Cambridge and London: Harvard University Press.

Slesnick, D. T. (1998): Empirical approaches to the measurement of welfare. *Journal of Economic Literature*, 36(4), 2108–2165.

Sørensen, P. B. (2005). Neutral taxation and shareholder income. *International Tax and Public Finance*, 12(6), 777–801.

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 labour supply: A discrete choice approach. *Journal of Human Resources*, 30(1), 63–88.

van Soest, A., M. Das, and X. Gong (2002). A structural labour supply model with flexible preferences. *Journal of Econometrics*, 107(1-2), 345–374.

Appendix

A.1. Empirical labour supply model for married couples, with estimation results

The modeling framework for two-person households is completely similar to the case for single-person households. Let $U\left(C,h_F,h_M,z\right)$ denote the utility function of a household, where h_F and h_M are hours of work for female and male and $z=\left(z_F,z_M\right)$ indexes the combination of jobs for the female and male in the household, respectively. As with one-person households, assume that $U\left(C,h_F,h_M,z\right)=u\left(C,h_F,h_M\right)+\varepsilon(z)$, with interpretation that is completely analogous to the case above. The budget constraint in this case can be written as

(A.1) $$C=f\left(h_F W_F,h_M W_M,I\right)$$

where W_F and W_M are the wage rates for female and male, respectively, and $f\left(.\right)$ is the function that transforms gross income to household disposable income. Let $\varphi\left(h_F,h_M \mid W_F,W_M,I\right)$ be the joint density of hours of work for female and male in the household, given wage rates and non-labour income. The empirical counterpart of this density is the fraction of couples where the husband works h_F hours and the wife works h_M hours, within the subpopulation of couples with wage rates and non-labor income equal to $\left(W_F,W_M,I\right)$. We assume furthermore that the offered hours, provided, H_F and H_F, are independent. For simplicity let

(A.2) $$\psi\left(h_F,h_M;W_F,W_M,I\right)=v\left(f\left(h_F W_F,h_M W_M,I\right),h_F,h_M\right).$$

Then, under exactly the same assumptions as for one-person households we obtain the conditional density of $\left(h_F,h_M\right)$, which, given that $h_M>0$, equals

(A.3) $$\varphi\left(h_F,h_M \mid W_F,W_M,I\right)=\frac{\psi\left(h_F,h_M;W_F,W_M,I\right)m_F\left(h_F\right)m_M\left(h_M\right)}{M},$$

for $h_F>0, h_M>0$, and

(A.4) $$\varphi\left(0,h_M \mid W_F,W_M,I\right)=\frac{\psi\left(0,h_M;W_F,W_M,I\right)m_M\left(h_M\right)}{M}$$

for $h_F=0$ and $h_M=0$, where

(A.5) $$M\left(W_F,W_M,I\right)=\sum_{y>0}\psi\left(0,y;W_F,W_M,I\right)m_M(y)+\sum_{x,y>0,}\psi\left(x,y\,|\,W_F,W_M,I\right)m_F(x)m_M(y).$$

The opportunity measure for females, θ_F, is assumed to depend on wages based solely on the amount of schooling. Specifically, we assume that

(A.6) $$\log\theta_F=f_{F1}+f_{F2}S,$$

where *S* is the length of education. Furthermore, we specify $v(\cdot)$ as having the form

(A.7) $$\log v\left(C,h_F,h_M\right)=\alpha_2\left(\frac{\left[10^{-4}\left(C-C_0\right)\right]^{\alpha_1}-1}{\alpha_1}\right)+\left(\frac{\left(L_F\right)^{\alpha_3}-1}{\alpha_3}\right)\beta_F$$
$$+\left(\frac{\left(L_M\right)^{\alpha_4}-1}{\alpha_4}\right)\beta_M+\alpha_{15}\left(\frac{\left(L_M\right)^{\alpha_4}-1}{\alpha_4}\right)\left(\frac{\left(L_F\right)^{\alpha_3}-1}{\alpha_3}\right)$$

where $\beta_F=\alpha_5+\alpha_6\log A_F+\alpha_7\left(\log A_F\right)^2+\alpha_8CU6+\alpha_9CO6$ and $\beta_M=\alpha_{10}+\alpha_{11}\log A_M+\alpha_{12}\left(\log A_M\right)^2+\alpha_{13}CU6+\alpha_{14}CO6$, C_0 is subsistence level, A_k, $k=F,M$, is the age for gender *k* divided by 10, *CU*6 and *CO*6 are the number of children under or equal to and over the age of six, respectively, *C* is given by the budget constraint in the same way as (2.2), L_k, $k=F,M$, represent leisure for gender *k*, with $L_k=1-h_k/3650$, and α_j, $j=1,2,...,15$, are unknown parameters. Note that we have subtracted from total annual hours a "subsistence" level, L_0 = 5,110 hours, which allows for sleep and rest. This corresponds to about 14 hours per day reserved for sleep and rest. We have chosen C_0 to be approximately NOK 40,000 $\sqrt{\text{N}}$, where *N* is the number of persons in the household. Disposable income, *C*, is measured as the sum of the annual household wage incomes after tax, household capital income after tax, and child allowances.

Table A.1. Estimates of wage equations, females and males, 1997

	Males		Females		Females (selection corrected)	
	Estimate	*t*-value	Estimate	*t*-value	Estimate	*t*-value
Constant	4.08	135.1	4.10	132	4.11	109
Experience in years/10	0.22	12.2	0.143	8.6	0.141	7.8
(Experience in years/10)^2	–0.03	–10.1	–0.022	–6.6	–0.022	–6.1
Education in years	0.044	26.9	0.0388	23.1	0.0386	19.7
Married	0.05	6.02	–0.022	–2.67	–0.21	–2.37
Log(P)					0.013	0.3
Variance of error term	0.3029		0.2755		0.2755	
No. of observations	5,448		5,074		5,074	
R^2	0.15		0.10		0.10	

Table A.2. Estimates of utility function parameters. Married couples, 1997. Box-Cox utility function.

	Parameter	Married couples	
		Estimate	SE
Preferences:			
Consumption			
Exponent	α_1	0.6643	0.054
Scale 10^{-4}	α_2	1.8411	0.352
Subsistence	C_0	$40{,}000\sqrt{N}$	
Female leisure			
Exponent	α_3	–0.8334	0.182
Constant	α_5	11.8387	1.888
Log(age/10)	α_6	–12.5285	1.945
Log(age/10) squared	α_7	5.2456	0.733
No. children under or equal to 6 years	α_8	0.9682	0.168
No. children over 6 years	α_9	0.5075	0.094
Male leisure			
Exponent	α_4	–1.8043	0.430
Constant	α_{10}	3.8929	1.112
Log(age/10)	α_{11}	–4.3054	1.142
Log(age/10) squared	α_{12}	1.6682	0.444
No. children under or equal to 6 years	α_{13}	0.0547	0.051
No. children over 6 years	α_{14}	0.0083	0.029
Leisure interaction	α_{15}	0.2047	0.147
Leisure subsistence	L_0	5,110	
The parameters θ_F ; $\log\theta_F = f_{F1} + f_{F2}S$			
Constant	f_{F1}	–3.5041	0.435
Education	f_{F2}	1.2389	0.366
Opportunity density of hours offered			
Male full-time peak		2.3769	0.086
Female full-time peak		1.4380	0.296
Male part-time peak		1.0960	0.063
Female part-time peak		0.5622	0.067
Number of observations		2,511	
Log likelihood		–5,706.5	
McFadden's ρ^2		0.44	

A.2. Properties of the indirect utility function

Given the distributional assumptions of the error terms the distribution of $\max_{z\in B(h)} \varepsilon(z,h)$ can easily be obtained. To realise this note that since $\varepsilon(z,h)$, z = 1,2,…, are i.i.d. with c.d.f. $\exp(-e^{-x})$ it follows that

$$P\left(\max_{z\in B(h)} \varepsilon(z,h) \le x\right) = P\left(\bigcap_{z\in B(h)} (\varepsilon(z,h) \le x)\right) = \prod_{z\in B(h)} P(\varepsilon(z,h) \le x)$$

$$= \prod_{z\in B(h)} \exp(-e^{-x}) = \exp(-\sum\nolimits_{z\in B(h)} e^{-x}) = \exp(-\theta g(h) e^{-x}).$$

Accordingly, if we let $\eta_h = \max_{z\in B(h)} \varepsilon(z,h) - \log(\theta g(h))$ we obtain from the equation above that

$$P(\eta_h \le x) = P\left(\max_{z\in B(h)} \varepsilon(z,h) \le x + \log(\theta g(h)\right) = \exp\{-\theta g(h) \exp(-x - \log(\theta g(h))\} = \exp(-e^{-x}),$$

which means that the error term η_h has the same distribution as $\varepsilon(z,h)$. Moreover, η_h and η_k are independent due to the fact that the choice sets $B(h)$ and $B(k)$ contain different jobs and the error terms $\varepsilon(z,h)$ and $\varepsilon(z',k)$ are independent when $z \ne z'$. We can therefore write the conditional indirect utility function in (2.2a,b) as

(A.8) $$V(h,y) = \log \mu + u(f(hw,y),h) + \log(\theta g(h)) + \eta_h$$

for $h > 0$ and

(A.9) $$V(0,y) = u(f(0,y),0) + \varepsilon(0).$$

A.3. Analytic formulae for the distribution of changes in welfare

In the random utility framework, the model formulation implies that *CV* and *EV* become random variables, and in general no closed-form formulae for welfare effects exist. This subsection provides a short review of the approach introduced by Dagsvik and Karlström (2005), giving formulae for the distribution and moments for *CV*. Note that a procedure for obtaining the same characteristics for *EV* is for all practical purposes identical. The pre-reform values are represented by (f^0, y), with (f^1, y) denoting the post-reform structure. We now need to introduce observed individual characteristics into the notation. Specifically, let the subscript *i* index the individual and let X_i be a vector of observed individual characteristics of individual *i*. It follows from the definition above that CV_i is determined by

(A.10) $$\max_{h\in D}(u(f^0(w_i h, y_i), h, X_i) + \eta_{ih}) = \max_{h\in D}(u(f^1(w_i h, y_i) + CV_i, h, X_i) + \eta_{ih}).$$

A key assumption here is that the error terms do not change from the ex-ante to the post-reform situation. Recall that $u(f^{t}(w_i h, y_i), h, X_i) + \eta_{ih}$ is a conditional indirect utility function in the sense that it is the utility of the most preferred job within the latent choice set $B(h)$. This means that our notion of compensating variation takes into account that the choices of the individual agents are constrained choices. Specifically, conditional on X_i, CV is a random variable that in general depends on all stochastic terms $\{\eta_{ih}, h \in D\}$. This is because the maximum of the left-hand side of (A.3) is not necessarily attained at the same discrete alternative as the maximum of the right-hand side of (A.3), except in special cases. Consequently, the random terms on each side do not cancel. Thus, the compensating measure defined above depends on the stochastic terms that represent unobserved preference heterogeneity as well as on $\theta g(h)$, which represents the value of the choice sets. We also assume that $u(f^{r}(hw_i, y_i) + x, h, X_i) = -\infty$ when $x \le a - f^{1}(hw_i, y_i) := c_i$ where *a* is some suitable small number. This ensures that the utility of income less than some suitable value becomes very small.

Let y_{ih} be the value that solves $u(f^{0}(w_i h, y), h, X_i) = u(f^{1}(w_i h, y_i) + y_{ih}, h, X_i)$, $\tilde{h}_i^0$ the ex-ante chosen hours of work and

$$K(x,h \mid X_i) = P(CV_i > x, \tilde{h}_i^0 = h \mid X_i).$$

In words, the latter expression is the joint probability that the compensating variation, given the initial utility level, is greater than some given real number *x* and the ex-ante chosen hours of work equals *h*. Dagsvik and Karlström (2005) demonstrate that the conditional distribution (1 minus the c.d.f.) of the expenditure function, given the individual characteristics, is given by

(A.11) $$K(x,h \mid X_i) = \frac{1\{x \ge \max(a, y_{ih})\}\exp(u(f^{0}(w_i h, y_i), h, X_i))}{\sum_{q\in D}\exp(\max(u(f^{0}(w_i q, y_i), q, X_i), u(f^{1}(w_i q, y_i) + x, q, X_i))},$$

where the notation $1\{x\}$ means that $1\{x\} = 1$ if *x* is true and zero otherwise. The distribution of the expenditure function in Eq. (3.2) can be used to find the first and second order moments of *CV* conditional on X_i. The conditional expected value can be computed as

(A.12) $$E(CV_i \mid X_i) = \int_{c_i}^{\infty} P(CV_i > x \mid X_i)dx + c_i \; = c_i + \sum_{h\in D}\int_{c_i}^{y_{ih}} K(x,h \mid X_i)dx$$

where $P(CV_i > c_i \mid X_i) = 1.$ Similarly, the conditional variance can be readily calculated. Standard numerical approaches can be applied to evaluate the indefinite integral in (A.12). Provided a representative micro-population *S* (say) is available, the corresponding unconditional expectation can be approximated by

(A.13) $$ECV_i \cong \frac{1}{N}\sum_{i\in S} E(CV_i \mid X_i)$$

where N is the micro-population size. We note that actual labour supply choices play no role in the calculation of the distribution function. Kornstad and Thoresen (2006) and Dagsvik, Locatelli, and Strøm (2009) apply this method to obtain the distribution of *CV*.

A.4. McFadden's simulation approach

McFadden (1999) describes a practical procedure for obtaining average estimates of *EV* and *CV*, based on simulations. Here we demonstrate how measures of *CV* are obtained, but corresponding formulae can be derived for *EV*. The starting point for obtaining estimates of *CV* by the simulation approach of McFadden (1999) can be described as

(A.14) $$\max_{h\in D}(u(f^0(w_i h, y_i), h, X_i) + \eta_i^k(h)) = \max_{h\in D}(u(f^1(w_i h, y_i) + CV_i, h, X_i) + \eta_i^k(h)),$$

for each random generated set, k, of standard Gumbel distributed error terms $\{\eta_i^k(h), h\in D, i\in S, k\in M\}$. At first glance it might seem to be numerically time-consuming to solve equation (A.14) directly, since it requires a maximisation algorithm. This may be so in the case where the random error terms are correlated, but not in our case where the error terms are independent. This means that a standard one-dimensional search algorithm can be applied.

In practice, many sets of draws are obtained for the random part of the utility function. The corresponding *CV* is then calculated by solving (A.14) for each set of draws. From these data the simulated distribution of *CV* can be computed as well as the corresponding mean and variance. This can be done in two ways: one way is to compute the mean *CV* for each individual by taking the mean across all individual-specific simulations and subsequently calculating the empirical distributions of these individual-specific means. A second way is to calculate the distribution using all the simulated *CV* values directly. The interpretation of the latter approach is that it accounts for unobserved heterogeneity within observationally identical groups of individuals.