CESifo Working Papers
7493
2019
January 2019

Marginal Compensated Effects in Discrete Labor Supply Models

John K. Dagsvik, Steinar Strøm, Marilena Locatelli

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7493*
Category 4: Labour Markets

Marginal Compensated Effects in Discrete Labor Supply Models

Abstract

This paper develops analytic results for marginal compensated effects of discrete labor supply models, including Slutsky equations. It matters, when evaluating marginal compensated effects in discrete choice labor supply models, whether one considers wage increase (right marginal effects) or wage decrease (left marginal effects). We show how the results obtained can be used to calculate the marginal cost of public funds in the context of discrete labor supply models. Subsequently, we use the empirical labor supply model of Dagsvik and Strøm (2006) to compute numerical compensated (Hicksian) and uncompensated marginal (Marshallian) effects resulting from wage changes. The mean Hicksian labor supply elasticities are larger than the Marshallian, but the difference is small.

JEL-Codes: J220, C510.

Keywords: Slutsky equations, discrete choice labor supply.

John K. Dagsvik
The Ragnar Frisch Centre of Economic Research / Oslo / Norway
John.Dagsvik@ssb.no

Steinar Strøm
Department of Economics
University of Oslo & The
Ragnar Frisch Centre of Economic Research
Oslo / Norway
steinar.strom@econ.uio.no

Marilena Locatelli
The Ragnar Frisch Centre of Economic Research / Oslo
& Department of Economics and Statistics
University of Turin / Italy
marilena.locatelli@unito.it

January 22, 2019
We acknowledge financial support of Research Grants from the University of Oslo and from the Ragnar Frisch Centre of Economic Research, Oslo.

1. Introduction

In microeconomic theory the Slutsky equation plays a fundamental role in the calculation of marginal compensated effects (Slutsky, 1915, Hicks, 1936, Varian, 1992). It allows one to compute the marginal compensated (Hicksian) labor supply effect (which is unobservable) from the corresponding Marshallian marginal labor supply effect. The compensated labor supply elasticity is one of the key parameters in the calculation of the deadweight loss of taxation and in the devising an optimal income tax policy: see Feldstein (1999) and Saez (2001). Another example is the calculation of the marginal cost of public funds, where compensated labor supply responses are widely used: see Jacobs (2018).

In traditional labor supply models, where the hours of work equation is usually given by a closed-form expression, it is straightforward to calculate the compensated marginal effects. In discrete labor supply models, the situation is different. Here the labor supply function cannot be expressed on closed form. Furthermore, it is stochastic and not differentiable, so one cannot compute marginal, individual effects in the usual way. However, one can compute marginal compensated effects of the corresponding probability distributions and expectations of labor supply functions. In other words, it is possible to compute marginal aggregate compensated labor supply effects.[4]

In this paper we shall demonstrate that aggregate Slutsky-type equations exist for discrete labor supply models and can be expressed in a convenient way. Subsequently, we propose a novel approach to compute marginal cost of public funds in the context of discrete labor supply models.

Since the mid-1990s empirical labor supply based on the theory of discrete choice and random utility representations has become increasingly popular. A major reason for this is that discrete choice labor supply models are much more practical than the conventional approach based on marginal calculus with a continuum of hours available: see Creedy and Kalb (2005) and the survey by Dagsvik et al. (2014). With the discrete choice approach, it is easy to deal with non-linear and non-convex economic budget constraints, and to apply rather general functional forms in utility representations.

In the literature, basically two versions of discrete models of labor supply have been developed. Van Soest (1995) proposed analyzing labor supply as a standard discrete choice problem. Dagsvik (1994) proposed a version of the discrete labor supply model framework, later denoted the job choice model, which was initially applied empirically by Dagsvik et al. (1988) and Dagsvik and Strøm (1992), with further extensions by Dagsvik and Strøm (2006) and Dagsvik and Jia (2016).[5] This last paper also provides an analysis of the identification problem in this kind of model.

[4] The first attempt to obtain Slutsky type of relations for discrete choice models seems to be Small and Rosen (1981). Their analysis is however of minor relevance for our approach. In the first place, their analysis is about discrete/continuous commodity demand. Unfortunately, their analysis is not fully correct, as we shall explain below.

[5] Some applications are based on the same framework: see Aaberge et al. (1995) and Aaberge et al. (1999).

In the job choice model, labor supply behavior is viewed as a job choice problem, where the set of feasible jobs is individual-specific, latent, and finite. The job choice model includes the standard discrete choice labor supply model as a special case. From a theoretical point of view, the standard discrete labor supply model is similar to the traditional textbook model, where the agent's labor supply decision is based on maximization of utility with respect to consumption and hours of work, subject to budget constraint. The only new aspect is that the set of feasible hours of work is assumed to be discrete. The job choice model, however, differs from this setup in that the worker is also allowed to have preferences over non-pecuniary attributes of jobs. That is, the agent is viewed as choosing from a finite and agent-specific set of available "packages" (jobs), where each job is characterized by job-specific hours of work and latent non-pecuniary attributes. This framework allows us to accommodate observed peaks in the hours of work distribution, which is interpreted as the result of more jobs with full-time and part-time hours of work being available than jobs with other hours of work schedules.

In this paper we have focused on analytic results. Alternatively, one could carry out Monte Carlo (MC) simulations of marginal compensated effects. However, the task of deriving analytic results is of major interest for several reasons. First, it can reveal how marginal compensated and uncompensated effects (such as Slutsky-type equations) are related: for example, in which cases the marginal compensated and uncompensated wage elasticities are equal. Second, the existence of practical analytic results may also facilitate numerical computation. Third, analytic results are usually more precise than results based on MC simulations, because they are not plagued by simulation errors. Fourth, analytic result can be used to distinguish directly between the impact of wage increase and wage decrease on labor supply. The analytic results on compensated marginal effects show some unexpected features. Specifically, the left and right marginal effects are in general different. That is, the marginal compensated effect resulting from a wage increase differs from the marginal compensated effect resulting from a wage decrease.

By applying the same approach as in Dagsvik and Karlström (2005) and Dagsvik (2018), we have obtained marginal effects and Slutsky equations in discrete job choice labor supply models in a way that is similar to the textbook case. Subsequently, these analytic results are applied to compute marginal compensated effects based on the estimated empirical model of Dagsvik and Strøm (2006). In these applications it turns out that with a wage increase the compensated and the uncompensated wage elasticities of the mean hours of work are rather close, with the former being slightly larger than the latter. The right and left compensated elasticities related to the intensive margin are very close, even when the number of discrete alternatives in the choice set is low. However, at the extensive margin there is a substantial difference between right and left compensated marginal effects.

Based on a discrete labor supply model estimated on Norwegian data for married women, we find that the labor supply elasticities, compensated and uncompensated, at the extensive and intensive margins vary substantially across observed covariates such as wage level, non-labor income, age, and number of children. The labor supply elasticities, compensated and uncompensated, are greatest for those with the lowest wages, those with the highest non-labor incomes, and those with many children. Thus, according to our estimates of compensated labor supply elasticities, the distortionary effect of taxation is stronger for individuals with low wages than for those with high wages. Because of the heterogeneity at the micro level, the aggregate labor supply elasticity is not a structural parameter.

Our concept of marginal cost of funds is based on the aggregate Compensating Variation measure, derived from the random expenditure function, and the aggregate compensated tax revenue. Our estimate is in the lower range of what others have found, based on quite different approaches.

The paper proceeds as follows. In Section 2 we present variants of discrete labor supply models, ranging from the conventional discrete labor supply model (Van Soest, 1995) to the multisectoral job choice model (Dagsvik and Strøm, 2006). In Section 3 we derive compensated marginal effects and Slutsky equations in these labor supply models, Section 4 discusses how the aggregate marginal cost of funds can be calculated in discrete choice models. Section 5 illustrates the results empirically based on the multisectoral job choice model in Dagsvik and Strøm (2006). Section 6 concludes.

2. Variants of discrete labor supply models

2.1. The conventional discrete labor supply model

We first describe the conventional discrete choice labor supply model (Van Soest, 1995). Let *C* and *h* denote consumption (disposable income) and hours of work respectively. Let

$$U(C,h) = u(C,h) + \varepsilon_h$$

be the agent's utility, where $u(C,h)$ is a positive deterministic function that is strictly increasing in *C* and strictly decreasing in *h*. Let *w* and *y* denote the agent's wage rate and non-labor income and $f(x,y)$ the function that transforms labor and non-labor income to income after tax. Thus,

$$C = f(hw, y).$$

Hours of work *h* belongs to a finite set *D*, which includes zero hours. The terms $\varepsilon_h,\ h \in D,$ are random variables that are supposed to account for unobserved heterogeneity in preferences across alternatives and agents, and they are assumed to be i.i.d. with c.d.f. $\exp(-e^{-x})$ (Gumbel c.d.f.). Note that the random variables depend on the choice of hours. Let $\varphi(h) = \varphi(h; w, y)$ denote the probability of supplying *h* hours given the wage rate and non-labor income (w, y) and let $v(h) = v(h; w, y) = u(f(hw, y), h).$ It follows from well-known results that the Marshallian probability is

$$\varphi(h) = \frac{\exp(u(f(hw,y),h))}{\sum_{x \in D} \exp(u(f(xw,y),x))} \tag{2.1}$$

for $h \in D$.

2.2. The job choice labor supply model

As mentioned in the introduction, the job choice model, developed by Dagsvik (1994) and further developed by Dagsvik and Strøm (2006 and Dagsvik and Jia (2016), allows us to account for latent restrictions in the labor market. Such restrictions may explain why the distribution of hours of work typically show peaks at full-time and part-time hours of work. Furthermore, the job choice model can also accommodate the fact that workers face different restrictions on the set of available (latent) jobs. In this model the household derives utility from household consumption, leisure, and non-pecuniary latent job attributes.

Let *z* = 1, 2..., be an indexation of the jobs and let *z* = 0 represent not working. The utility function is assumed to have the form

(2.2) $$U(C,h,z) = u(C,h) + \zeta + \varepsilon(z)$$

for *h* > 0 and $\zeta = 0$ when *h* = 0. The terms $\{\varepsilon(z)\}$ are sector- and job-specific random taste shifters. The taste shifters $\{\varepsilon(z)\}$ are assumed to be i.i.d. across jobs and agents, with c.d.f. $\exp(-e^{-x})$, for real *x*. The taste shifters account for unobserved individual characteristics and unobserved job-specific attributes. The term ζ represents the mean preference for the mean non-pecuniary value of working. Let *B*(*h*) be the set of jobs with hours of work *h* that are available to the agent. The sets *B*(*h*), $h \in D$, are individual-specific and latent. Moreover, let $\tilde{\theta}$ be the number of jobs in *B*(*h*) and *g*(*h*) the proportion of jobs in *B*(*h*) with hours of work *h*. Thus, $\tilde{\theta} g(h)$ is the number of jobs with hours of work *h* in the latent set $B(h)$. From (2.2) it follows that the highest utility the agent can attain, bearing in mind that hours of work are equal to *h*, is given by

(2.3) $$V(h,w,y) := \max_{z \in B(h)} U(f(hw,y),h,z)$$

$$= u(f(hw,y),h) + \max_{z \in B(h)} \varepsilon(z) = u(f(hw,y),h) + \log(\theta g(h)) + \eta(h)$$

where $\theta = \tilde{\theta} \exp(\zeta)$ and

$$\eta(h) = \max_{z \in B(h)} \varepsilon(z) - \log(\theta g(h)).$$

It follows that $\eta(h)$ has the same c.d.f. as $\varepsilon(1)$ because

$$P(\eta(h) \le x) = P\left(\max_{z \in B(h)} \varepsilon(z) - \log(\theta g(h)) \le x\right) = \prod_{z \in B(h)} P(\varepsilon(z) \le x + \log(\theta g(h)))$$

$$= \exp\{-\sum_{z \in B(h)} \exp(-x - \log(\theta g(h))) = \exp(-\theta g(h) \exp(-x - \log(\theta g(h)))\} = \exp(-e^{-x}).$$

Thus, formally, the utility maximization (with respect to hours of work) in the presence of these types of latent constraints can be achieved from the corresponding unconstrained case by modifying the structural part of the utility function by replacing $u(f(hw,y),h)$ with $u(f(hw,y),h)+\log(\theta g(h))$ for $h>0$, whereas $u(f(0,y),0)$ remains unchanged.

Let $\varphi(h)$ be the Marshallian probability of choosing hours of work h (for a utility maximizing agent). From (2.3) it follows immediately from the theory of discrete choice that

$$\text{(2.4)} \quad \varphi(h) = P(V(h,w,y) = \max_{x\in D\cup\{0\}} V(x,w,y)) = \frac{\exp(u(f(hw,y),h))\theta g(h)}{\exp(u(f(0,y),0)) + \theta \sum_{x\in D\setminus\{0\}} \exp(u(f(xw,y),x))g(x)}$$

for $h>0$. For $h = 0$, $\varphi(0)$ is obtained from (2.4) by replacing the numerator by $\exp(u(f(0, y), 0))$.

From (2.3) and (2.4) we note that the job choice model has, formally, the same mathematical form as the conventional discrete choice model with the systematic part of the utility function equal to $u(f(hw,y),h)+\log(g(h)\theta)$. Dagsvik and Jia (2016) have discussed identification of u, θ and $g(h)$. Note that the model considered in Section 2.1 follows as a special case when $\theta g(h)=1$.

Let

$$\text{(2.5)} \qquad \gamma(h) = \frac{f_1'(hw,y)h}{f_2'(hw,y)}.$$

The uncompensated (Marshallian) marginal wage effect in this model can be readily expressed as

$$\text{(2.6)} \qquad \frac{\partial\varphi(h)}{\partial w} = \varphi(h)\left(\frac{\partial v(h)}{\partial w} - \sum_{x\in D}\frac{\partial v(x)}{\partial w}\varphi(x)\right) = \varphi(h)\left(\frac{\partial v(h)}{\partial y}\gamma(h) - \sum_{x\in D}\frac{\partial v(x)}{\partial y}\gamma(x)\varphi(x)\right).$$

2.3. The multisectoral job choice model

This section outlines the job choice model with several sectors. An empirical two-sector version of the job choice model was developed by Dagsvik and Strøm (2006). It was applied to conduct welfare analysis in Dagsvik et al. (2009). Let w_k denote the wage the agent receives when working in sector k, k = 1, 2,…The budget constraint when working in sector k is given by

$$C_k = f(hw_k, y)$$

and the utility function is assumed to have the structure

$$U_j(C_j,h,z) = u(C_j,h) + \zeta_j + \varepsilon_j(z)$$

where $u(C_j,h)$ is a deterministic term and ζ_j is a parameter that represents the average preference for sector j-specific tasks. For the non-working alternative, $\zeta_0 = 0$. Random taste shifters $\varepsilon_j(z)$ are i.i.d. across jobs and sectors and follow the extreme value distribution. Let $\varphi_j(h)$ be the Marshallian probability of choosing a job in sector j with hours of work h (for a utility maximizing agent). Let $\theta_j = \tilde{\theta}_j \exp(-\zeta_j)$ where $\tilde{\theta}_j$ denotes the total number of jobs available to the worker in sector j. This implies that θ_j is a measure of the total number of jobs available to the worker in sector *j* weighted by the attractiveness $(\exp(\zeta_j))$ of the sector. Let

$$v_j(h;w_j,y) = v_j(h,y) = u(f(hw_j,y),h) + \log(g_j(h)\theta_j)$$

and

$$v(0,y) = u(f(0,y),0).$$

Similarly to the previous section it now follows that

$$V_j(h,w_j,y) := \max_{z \in B_j(h)} U_j(f(hw_j,y),h,z) = v_j(h;w_j,y) + \eta_j(h)$$

where $B_j(h)$ is the set of jobs with hours of work *h* that are available to the agent. With this notation we can express the multisectoral choice probabilities in a compact way as

(2.7)
$$\varphi_j(h) = \frac{\exp(v_j(h,y))}{\exp(v(0,y)) + \sum_r \sum_{x \in D\setminus\{0\}} \exp(v_r(x,y))}$$

and

(2.8)
$$\varphi_j(0) = \frac{\exp(v_j(0,y))}{\exp(v(0,y)) + \sum_r \sum_{x \in D\setminus\{0\}} \exp(v_r(x,y))}.$$

We realize that the models considered in Sections 2.1 and 2.2 become, formally, special cases of the model in (2.7) and (2.8). This is convenient for our subsequent analysis because it allows a unified treatment.

In the presentation above it is not discussed how choice sets of jobs are generated. Dagsvik (2000), Menzel (2015) and Dagsvik and Jia (2018) have demonstrated that the job choice model can in fact be viewed as an equilibrium matching model with non-transferable preferences.

3. Compensated effects and Slutsky equations

In the traditional textbook case, where the commodity space is a continuum, the substitution effect can be visualized as a move along an indifference curve. To calculate the marginal compensated effects from a change in the wage rate in the textbook case, one can apply the Slutsky equation. To review the Slutsky equation, let $h(w, y)$ denote the Marshallian labor supply of hours of work, as a function of the wage rate and non-labor income (w,y), and let $h^H(w,u)$ denote the corresponding Hicksian labor supply function, where *u* is the utility level. At optimum, $h^H(w,u) = h(w,e(w,u))$, where $e(w,u)$ is the expenditure function needed to keep utility at the level *u*. Using that at optimum non-wage income *y* is equal to $e(w,u)$, and applying Shepard's lemma the Slutsky equation follows by differentiating through the expression above with respect to *w*, which yields

$$\frac{\partial h(w,y)}{\partial w}=\frac{\partial h^{H}(w,u)}{\partial w}-h(w,y)\frac{\partial h(w,y)}{\partial y}$$

This equation allows one to compute the marginal compensated Hicksian marginal labor supply effect (which is unobservable) from the corresponding marginal Marshallian labor supply effects. In traditional labor supply models, where the hours of work equation is usually given by a closed-form expression, it is straightforward to calculate these compensated marginal effects. In our case the labor supply function is stochastic and cannot be expressed on closed form. Therefore, another approach is called for. Our approach is based on Dagsvik and Karlström (2005) and aims at obtaining analytic results for the distribution and expectation of labor supply. To the best of our knowledge, Slutsky equations for discrete labor supply models have not been obtained previously. Slutsky equations for general discrete choice models are discussed by Dagsvik (2019). However, the results obtained in this paper are not special cases of Dagsvik (2019).

We shall now consider marginal effects. We take a setting where there is a change in the wage rate from the initial ex ante value w to the ex post value $\tilde{w}$. For simplicity we start with only one sector. Let $Q^{H}(x,h)=Q^{H}(x,h;w,y,\tilde{w})$ be the joint compensated probability of choosing *x* hours of work ex ante and *h* hours of work ex post, *w* and $\tilde{w}$ the wage rate ex ante and ex post respectively, and *y* the ex ante non-labor income. Thus, here the ex ante and ex post utility levels are equal. Dagsvik and Karlström (2005) proved the following result:

Theorem 1

Assume a random utility model $U(h)=v(h;w,y)+\varepsilon(h)$ *where* $\{\varepsilon(h),h\in D\}$ *are independent and standard Gumbel-distributed and let* y_h *be defined by* $v(h;w,y)=v(h;\tilde{w},y_h(\tilde{w}))$. *Then*

(3.1) $$Q^{H}(x,h)=\int_{y_h(\tilde{w})}^{y_x(\tilde{w})}\frac{\exp(v(x;w,y)+v(h;\tilde{w},z))v(h;\tilde{w},dz)}{M(z)^{2}}$$

when $y_x(\tilde{w})\geq y_h(\tilde{w})$, $Q^{H}(x,h)=0$ *when* $y_x(\tilde{w})<y_h(\tilde{w})$, *and*

(3.2) $$Q^{H}(h,h)=\frac{\exp(v(h;w,y)}{M(y_h(\tilde{w}))}$$

where

$$M(z)=\sum_{x\in D}\exp(\max(v(x;w,y),v(x;\tilde{w},z))$$

Proceeding from this theorem, let $P^{H}(h;w,\tilde{w},y)$ be the compensated (Hicksian) probability of choosing *h* hours of work ex post, given that the indirect utility is kept fixed. It follows that

(3.3) $$P^{H}(h;w,\tilde{w},y)=\sum\nolimits_{x\in D}Q^{H}(x,h).$$

We wish to compute the compensated marginal effect of the choice probability of working *h* hours. By this we mean

$$\frac{\partial \varphi^{H}(h)}{\partial w} = \lim_{\Delta w \to 0} \frac{P^{H}(h; w, w+\Delta w, y) - \varphi(h; w, y)}{\Delta w}$$

where $\Delta w = \tilde{w} - w.$ However, it turns out that the above derivative does not always exist. We therefore need to introduce the left and right derivatives, defined in the usual way as

(3.4) $$\frac{\partial^{+} \varphi^{H}(h)}{\partial w} = \lim_{\Delta w \downarrow 0} \frac{P^{H}(h; w, w+\Delta w, y) - \varphi(h; w, y)}{\Delta w}.$$

and

(3.5) $$\frac{\partial^{-} \varphi^{H}(h)}{\partial w} = \lim_{\Delta w \uparrow 0} \frac{P^{H}(h; w, w+\Delta w, y) - \varphi(h; w, y)}{\Delta w}.$$

In (3.4) $\Delta w > 0$ and approaches zero from above, whereas in (3.5) $\Delta w < 0$ and approaches zero from below. The formula in (3.4) is the right derivative of the Hicksian probability of working *h* hours with respect to the wage rate. This formula is relevant for computing the compensated marginal effect of an *increase* in the wage rate. The formula in (3.5) is the corresponding left derivative, which is relevant for computing the compensated marginal effect of a *decrease* in the wage rate. Recall that the derivative $\partial \varphi^{H}(h) / \partial w$ exists only if $\partial^{+} \varphi^{H}(h) / \partial w = \partial^{-} \varphi^{H}(h) / \partial w.$

Let *Z* be a real number and define $Z_{+} = \max(Z, 0).$ We then have the following result:

Theorem 2

Under the assumptions of Theorem 1 the compensated marginal effects in the conventional discrete labor supply model are given by

$$\frac{\partial^{+} \varphi^{H}(h)}{\partial w} = \frac{\partial v(h)}{\partial y} \varphi(h) \sum_{x \in D} \varphi(x) (\gamma(h) - \gamma(x))_{+} - \varphi(h) \sum_{x \in D} \varphi(x) \frac{\partial v(x)}{\partial y} (\gamma(x) - \gamma(h))_{+}$$

and

$$\frac{\partial^{-} \varphi^{H}(h)}{\partial w} = \varphi(h) \sum_{x \in D} \varphi(x) \frac{\partial v(x)}{\partial y} (\gamma(h) - \gamma(x))_{+} - \varphi(h) \frac{\partial v(h)}{\partial y} \sum_{x \in D} \varphi(x) (\gamma(x) - \gamma(h))_{+}$$

for $h \in D.$

Theorem 2 is a special case of Theorem 3 given in Section 3.3 and proved in Appendix B. In contrast to the traditional case, the formulas given in Theorem 2 express marginal aggregate compensated effects. Corollary 1 below follows readily from (2.6) and Theorem 2.

At first glance it would seem possible to apply the same approach as Small and Rosen (1981) to obtain marginal compensated effects. However, their analysis is not fully correct. Although the formulas for aggregate marginal compensated probabilities given on page 117 in their paper appear to be correct their relations to choice probabilities derived from stochastic utility specification, such as eq. (5.1) in their paper, is not made clear. In fact, the marginal compensated probabilities given on page 117 are not necessarily the same as the corresponding marginal compensated effects derived from random utility specifications such as in (5.1). Consequently, their formulas on page 117 on the Slutsky relations cannot be applied in our context. Further discussion on the Small and Rosen (1981) setup is given in Dagsvik (2019).

Corollary 1

Under the assumptions of Theorem 1 the corresponding right Slutsky equation for the labor supply probabilities is given by

$$\frac{\partial\varphi(h)}{\partial w}=\frac{\partial^{+}\varphi^{H}(h)}{\partial w}-\varphi(h)\sum_{x\in D}\left(\frac{\partial v(x)}{\partial y}-\frac{\partial v(h)}{\partial y}\right)\varphi(x)\min\left(\gamma(h),\gamma(x)\right)$$

and the corresponding left Slutsky equation is given by

$$\frac{\partial\varphi(h)}{\partial w}=\frac{\partial^{-}\varphi^{H}(h)}{\partial w}-\varphi(h)\sum_{x\in D}\left(\frac{\partial v(x)}{\partial y}-\frac{\partial v(h)}{\partial y}\right)\varphi(x)\max\left(\gamma(h),\gamma(x)\right).$$

From Corollary 1 and (2.5) it follows immediately that the right and left Slutsky equations at the extensive margin are

$$\frac{\partial\varphi(0)}{\partial w}=\frac{\partial^{+}\varphi^{H}(0)}{\partial w}$$

and

$$\frac{\partial\varphi(0)}{\partial w}=\frac{\partial^{-}\varphi^{H}(0)}{\partial w}-\varphi(0)\sum_{x\in D}\left(\frac{\partial v(x)}{\partial y}-\frac{\partial v(0)}{\partial y}\right)\gamma(x)\varphi(x).$$

The relations given in Corollary 1 share some similarities with the traditional Slutsky equation for continuous choice, but they also differ in essential ways. As for the right Slutsky relation and for $h>0$, the aggregate income effect is given by

$$-\varphi(h)\sum_{x\in D}\left(\frac{\partial v(h)}{\partial y}-\frac{\partial v(x)}{\partial y}\right)\varphi(x)\min\left(\gamma(h),\gamma(x)\right),$$

which can be positive or negative, and even equal to zero. Disregarding for the moment the case of backward-bending labor supply, the point is that when (say) the wage increases the compensated as well as the uncompensated probabilities of working few hours will decrease and the respective probabilities for working many hours will increase. Note that there exist hours of work greater than zero where the difference between the wage derivatives of the compensated and uncompensated choice probabilities equals zero. Moreover, we can see that at the extensive margin there is no income effect in the right derivative case, but that there is an income effect in the left derivative case.

From Corollary 1 it also follows that

(3.6) $$\frac{\partial^- \varphi^H(h)}{\partial w} - \frac{\partial^+ \varphi^H(h)}{\partial w} = \varphi(h)\sum_{x\in D} |\gamma(h)-\gamma(x)| \left(\frac{\partial v(h)}{\partial y} - \frac{\partial v(x)}{\partial y}\right)\varphi(x).$$

The expression in (3.6) shows that the sign of the difference between the right and left marginal compensating effects may be positive or negative, and also zero. From Corollary 1 it also follows that if the deterministic part of the utility function $u(f(wh, y), h)$ is linear in after-tax income, and wage income and non-labor income are taxed separately, then the compensated marginal effects equal the corresponding uncompensated effects and accordingly the income effects in the Slutsky equation disappears.

In contrast to the standard textbook case, in the discrete labor supply setting the wage and hence the income derivatives of the deterministic part of the utility function in all alternatives have to be calculated. The reason is that a wage change affects all alternatives in the choice set, not only the alternative in focus.

To gain more intuition, it might be instructive to consider the marginal compensated effects in the following binary case where the choice set consists of two alternatives, working *h* hours and not working. As above the wage rate increases from *w* to $\tilde{w}$, ceteris paribus. To this end, let Y be the (random) expenditure function defined by

$$\max_{h\in D}(v(h;w,y)+\varepsilon_h) = \max_{h\in D}(v(h;\tilde{w},Y)+\varepsilon_h).$$

Then $Q^H(h,0)=0.$ Note that when the ex ante and compensated ex post choices are "not working" the ex post and ex ante utilities are equal, so (3.2) implies that

$$Q^H(0,0) = P(\max(v(h;\tilde{w},y), v(h;w,y)) + \varepsilon_h < v(0;y)+\varepsilon_0) = P(v(h;\tilde{w},y)+\varepsilon_h < v(0;y)+\varepsilon_0).$$

Thus, it follows that the compensated wage effect equals

$$P^H(0;w,\tilde{w},y) - \varphi(0) = Q^H(h,0) + Q^H(0,0) - \varphi(0)$$

$$= P(v(h;\tilde{w},y)+\varepsilon_h < v(0;y)+\varepsilon_0) - P(v(h;w,y)+\varepsilon_h < v(0;y)+\varepsilon_0).$$

We realize that the last expression is the difference between two uncompensated choice probabilities, namely the probability of not working when the wage equals $\tilde{w}$ minus the probability of not working when the wage equals *w*. Accordingly, the right marginal compensated wage effect must be equal to the corresponding marginal uncompensated effect.

Consider next the case where $\tilde{w} < w.$ Here

$$Q^H(0,0) = P(\max(v(h;\tilde{w},y), v(h;w,y)) + \varepsilon_h < v(0;y)+\varepsilon_0)$$

$$= P(v(h;w,y)) + \varepsilon_h < v(0;y)+\varepsilon_0) = \varphi(0).$$

Furthermore, in this case we realize that $Y = y_h > y$ and

$$Q^H(h,h) = P(v(0;y)+\varepsilon_0 < v(h;w,y)+\varepsilon_h = v(h;\tilde{w},Y)+\varepsilon_h > v(0;Y)+\varepsilon_0)$$

$$= P(v(0;y)+\varepsilon_0 < v(h;w,y)+\varepsilon_h > v(0;Y)+\varepsilon_0) 1\{v(h;w,y) = v(h;\tilde{w},Y)\}$$

$$= P(v(h;w,y)+\varepsilon_h > \max(v(0;y), v(0;y_h)+\varepsilon_0)) = P(v(h;w,y)+\varepsilon_h > v(0;y_h)+\varepsilon_0).$$

Consequently, the compensated marginal wage effect in this case equals

$$P^H(0;w,\tilde{w},y)-\varphi(0)=Q^H(h,0)+Q^H(0,0)-\varphi(0)=Q^H(h,0)=\varphi(h)-Q^H(h,h)$$
$$=P(v(h;w,y)+\varepsilon_h>v(0;y)+\varepsilon_0)-P(v(h;w,y)+\varepsilon_h>v(0;y_h)+\varepsilon_0)$$
$$=P(v(h;w,y)+\varepsilon_h<v(0;y_h)+\varepsilon_0)-P(v(h;w,y)+\varepsilon_h<v(0;y)+\varepsilon_0)=1-Q^H(h,h).$$

We note that the situation in this case is different from that in the former case because the first expression in the difference above cannot be interpreted as an uncompensated choice probability. Therefore, the corresponding left marginal compensated wage effect differs from the corresponding marginal uncompensated wage effect.

Corollary 2

Let $\tilde{h}$ and $\tilde{h}'$ be independent draws from the labor supply p.d.f. $\varphi(h)$. Then the right and left Slutsky equations for aggregate hours of work are given by

$$\frac{\partial E\tilde{h}}{\partial w}=\frac{\partial^+ E\tilde{h}^H}{\partial w}-Cov\left(\tilde{h},-\frac{\partial v(\tilde{h})}{\partial y}\min\left(\gamma(\tilde{h}),\gamma(\tilde{h}')\right)\right)+Cov\left(\tilde{h}',-\frac{\partial v(\tilde{h})}{\partial y}\min\left(\gamma(\tilde{h}),\gamma(\tilde{h}')\right)\right)$$

and

$$\frac{\partial E\tilde{h}}{\partial w}=\frac{\partial^- E\tilde{h}^H}{\partial w}-Cov\left(\tilde{h},-\frac{\partial v(\tilde{h})}{\partial y}\max\left(\gamma(\tilde{h}),\gamma(\tilde{h}')\right)\right)+Cov\left(\tilde{h}',-\frac{\partial v(\tilde{h})}{\partial y}\max\left(\gamma(\tilde{h}),\gamma(\tilde{h}')\right)\right)$$

Furthermore,

$$\frac{\partial E\tilde{h}}{\partial w}\le\frac{\partial^+ E\tilde{h}^H}{\partial w}\qquad and \qquad \frac{\partial E\tilde{h}}{\partial w}\le\frac{\partial^- E\tilde{h}^H}{\partial w}.$$

The left and right Slutsky equations in Corollary 2 follow readily from Corollary 1. The proof of the inequalities in Corollary 2 goes as follows. Note, first, that $-\partial v(h)/\partial y$ and $\gamma(h)$ are increasing functions of *h* and, second, that the correlation between $\tilde{h}$ and $-\min(\gamma(\tilde{h}),\gamma(\tilde{h}'))\partial v(\tilde{h})/\partial y$ is stronger than the correlation between $\tilde{h}'$ and $-\partial v(\tilde{h})/\partial y\min(\gamma(\tilde{h}),\gamma(\tilde{h}'))$ because $\tilde{h}$ and $\tilde{h}'$ are independent. Hence, the income effect

$$-\left[Cov\left(\tilde{h},-\frac{\partial v(\tilde{h})}{\partial y}\min\left(\gamma(\tilde{h}),\gamma(\tilde{h}')\right)\right)+Cov\left(\tilde{h}',-\frac{\partial v(\tilde{h})}{\partial y}\min\left(\gamma(\tilde{h}),\gamma(\tilde{h}')\right)\right)\right]<0$$

from which the inequalities in Corollary 2 follow. This feature is analogous to the standard textbook case given that leisure is a normal good. We would expect the difference between the two covariances in most cases to be small and so we would expect the income effect above to be small. The implication is that we would expect the uncompensated and the compensated mean hours wage elasticities to be rather similar.

It is also of interest to note that even when the model is a continuous multinomial logit model, as in Dagsvik (1994), the difference between the right and left marginal wage effects does not disappear.

The Slutsky equations for the multisectoral job choice model are analogous to the corresponding one-sector case. They are given in the next theorem.

Theorem 3

Under the assumption of the multisectoral discrete job choice model the right and left Slutsky equations for the labor supply probabilities are given by

$$\frac{\partial \varphi_j(h)}{\partial w_j} = \frac{\partial^+ \varphi_j^H(h)}{\partial w_j} - \varphi_j(h) \sum_{x \in D\setminus\{0\}} \left(\frac{\partial v_j(x)}{\partial y} - \frac{\partial v_j(h)}{\partial y} \right) \varphi_j(x) \min\left(\gamma_j(h), \gamma_j(x)\right),$$

$$\frac{\partial \varphi_j(h)}{\partial w_j} = \frac{\partial^- \varphi_j^H(h)}{\partial w_j} - \varphi_j(h) \sum_{x \in D\setminus\{0\}} \left(\frac{\partial v_j(x)}{\partial y} - \frac{\partial v_j(h)}{\partial y} \right) \varphi_j(x) \max\left(\gamma_j(h), \gamma_j(x)\right)$$

$$+(1-\varphi_j)\varphi_j(h)\gamma_j(h)\frac{\partial v_j(h)}{\partial y} - \varphi_j(h) \sum_{r \neq j} \sum_{x \in D\setminus\{0\}} \frac{\partial v_r(x)}{\partial y} \gamma_j(h) \varphi_r(x) - \gamma_j(h)\varphi_j(h)\varphi(0)\frac{\partial v(0)}{\partial y}$$

for $j > 0$,

$$\frac{\partial^+ \varphi_j^H(h)}{\partial w_k} = \frac{\partial \varphi_j(h)}{\partial w_k}, \qquad \frac{\partial \varphi_j(h)}{\partial w_k} = \frac{\partial^- \varphi_j^H(h)}{\partial w_k} - \varphi_j(h) \sum_{x \in D\setminus\{0\}} \left(\frac{\partial v_k(x)}{\partial y} - \frac{\partial v_j(h)}{\partial y} \right) \varphi_k(x)\gamma_k(x)$$

for $j \neq k,\ j,k > 0$,

$$\frac{\partial^+ \varphi^H(0)}{\partial w_j} = \frac{\partial \varphi(0)}{\partial w_j}$$

and

$$\frac{\partial \varphi(0)}{\partial w_j} = \frac{\partial^- \varphi^H(0)}{\partial w_j} - \varphi(0) \sum_{x \in D\setminus\{0\}} \left(\frac{\partial v_j(x)}{\partial y} - \frac{\partial v(0)}{\partial y} \right) \gamma_j(x)\varphi_j(x)$$

where $v_j(h) = u_j(f(w_j h, y), h) + \log(\theta_j g_j(h))$ *and*

$$\gamma_j(h) = \frac{\partial v_j(h)/\partial w_j}{\partial v_j(h)/\partial y} = \frac{f_1'(w_j h, y)h}{f_2'(w_j h, y)}.$$

The proof of Theorem 3 is given in Appendix B. From Theorem 3 we observe that

(3.7)

$$\frac{\partial^- \varphi_j^H(h)}{\partial w_j} - \frac{\partial^+ \varphi_j^H(h)}{\partial w_j} = -(1-\varphi_j)\varphi_j(h)\gamma_j(h)\frac{\partial v_j(h)}{\partial y} + \varphi_j(h)\gamma_j(h)\sum_{r \neq j} \sum_{x \in D\setminus\{0\}} \frac{\partial v_r(x)}{\partial y}\varphi_r(x) + \varphi(0)\frac{\partial v(0)}{\partial y}$$

$$+\varphi_j(h)\sum_{r \neq j} \sum_{x \in D\setminus\{0\}} \varphi_r(x)\left(\frac{\partial v_j(x)}{\partial y} - \frac{\partial v_j(h)}{\partial y} \right) | \gamma_j(h) - \gamma_j(x) |.$$

From (3.7) we note that the sign of the difference between the right and left marginal compensating effects might happen to be both positive and negative. We observe that in the case of right marginal cross effect the marginal compensated and marginal uncompensated effects are equal.

Above, we have seen that the left and right marginal effects differ substantially in the sense that the algebraic expressions are quite different. But it remains to be seen to what extent the left and right marginal effects differ numerically. This issue will be examined in Section 5.

It is immediately clear that Theorem 3 is valid also for the conventional discrete choice model (Van Soest, 1995), extended to the multisectoral setting. Further detailed results relating to sectorial marginal effects are given in Appendix A. Here we show only some of the results that relate to the right marginal effects.

Let φ_j denote the Marshallian probability of working in sector *j*. It follows readily from Theorem 3 by adding the choice probabilities with respect to hours of work that the marginal effects related to the sectoral choice probabilities are given by the next corollary.

Corollary 3

Under the assumptions of Theorem 3 we have the following right and left Slutsky equations at the sectoral extensive margins:

$$\frac{\partial\varphi_j}{\partial w_j}=\frac{\partial^+\varphi_j^H}{\partial w_j},$$

$$\frac{\partial\varphi_j}{\partial w_j}=\frac{\partial^-\varphi_j^H}{\partial w_j}-\sum_{h\in D\backslash\{0\}}\varphi_j(h)\gamma_j(h)\sum_{r\neq j}\sum_{x\in D\backslash\{0\}}\frac{\partial v_r(x)}{\partial y}\varphi_r(x)-\varphi(0)\frac{\partial v(0)}{\partial y}\sum_{h\in D\backslash\{0\}}\varphi_j(h)\gamma_j(h)$$

$$+(1-\varphi_j)\sum_{x\in D\backslash\{0\}}\gamma_j(x)\frac{\partial v_j(x)}{\partial y}\varphi_j(x),$$

$$\frac{\partial\varphi_j}{\partial w_k}=\frac{\partial^+\varphi_j^H}{\partial w_k}\quad \textit{and}$$

$$\frac{\partial\varphi_j}{\partial w_k}=\frac{\partial^-\varphi_j^H}{\partial w_k}-\varphi_j\sum_{x\in D\backslash\{0\}}\frac{\partial v_k(x)}{\partial y}\gamma_k(x)\varphi_k(x)+\sum_{h\in D\backslash\{0\}}\frac{\partial v_j(h)}{\partial y}\varphi_j(h)\sum_{x\in D\backslash\{0\}}\gamma_k(x)\varphi_k(x)$$

for $k\neq j,\ j,k>0.$

The Slutsky equation for mean hours in the case of a wage increase is

$$\frac{\partial E\tilde{h}_j}{\partial w_j}=\frac{\partial^+E\tilde{h}_j^H}{\partial w_j}-Cov\left(\tilde{h}_j,-\frac{\partial v_j(\tilde{h}_j)}{\partial y}\min\left(\gamma_j(\tilde{h}_j),\gamma_j(\tilde{h}_j')\right)\right)-Cov\left(\tilde{h}_j',-\frac{\partial v_j(\tilde{h}_j)}{\partial y}\min\left(\gamma_j(\tilde{h}_j),\gamma_j(\tilde{h}_j')\right)\right),$$

The corresponding Slutsky equation in the case of a wage decrease is given Corollary A1 in Appendix A, where it is also shown that the income effects are negative, as in the standard text book case:

$$\frac{\partial^+ E\tilde{h}_j^H}{\partial w_j} - \frac{\partial E\tilde{h}_j}{\partial w_j} > 0.$$

Remember that although the formulas above are based on a multinomial logit formulation they can easily be extended to the mixed logit case by allowing some parameters to be random effects.

4. Marginal cost of public funds

4.1. Approaches in the literature

Pigou (1947), Harberger (1964), and Browning (1976, 1987) introduced the concept of (compensated) marginal cost of public funds as a measure of the cost of a marginal change in public revenue, defined as the reduction in consumers' surplus relative to the increase in tax revenue. If revenue is redistributed to the consumers as a lump-sum tax, income effects of the tax change are neutralized and the marginal cost of public funds relates only to the distortionary effect of the tax change. In order to calculate the compensated marginal cost of public funds one needs the corresponding compensated labor supply elasticities. Stiglitz and Dasgupta (1971) and Atkinson and Stern (1974) applied the corresponding uncompensated marginal cost of public funds, which means that the income effects of the marginal tax change in question are not neutralized through lump-sum transfers. In this case the marginal cost of funds is evaluated using the uncompensated labor supply elasticities.

Marginal cost of funds is widely used in cost-benefit analysis. In Norway, the Ministry of Finance has set the compensated marginal costs of funds at 1.2, meaning that if a public investment is financed through taxation the relevant cost to be used in the cost-benefit analysis is 1.2 times the cost of the investment. Kleven and Kreiner (2006) report values for the marginal costs of public funds when accounting for labor supply responses at both the extensive and the intensive margins. They report values based on uncompensated as well as compensated labor supply elasticities at the extensive and intensive margins for five European countries (Denmark, France, Germany, Italy, and the UK). The range of the marginal cost of public funds estimates is 1.26–2.2.

It may be interesting to review how measures of marginal compensated cost of public funds proposed in the literature relates to the one we propose below. To review this approach briefly, let $V(y)$ denote the indirect utility of the representative agent (RA) as a function of the ex ante non-labor income *y*. Let $\partial R^H(u)/\partial t$ denote the marginal compensated revenue at utility level *u*. The corresponding marginal compensated cost of public fund at utility level *u* is defined by

(4.1) $$MCF^H = -\frac{\frac{\partial V(y)}{\partial t} / \frac{\partial V(y)}{\partial y}}{\frac{\partial R^H(u)}{\partial t}\Big|_{u=V(y)}}.$$

We shall now demonstrate that the measure defined in (4.1) is in fact similar to the one shown in Jacobs (2018). Let $e(u)$ be the expenditure function at utility level *u*.

Since $V(e(u)) = u$ differentiation with respect to *t*, keeping utility level *u* fixed yields

(4.2) $$-\left[\frac{\partial V(y)}{\partial t} / \frac{\partial V(y)}{\partial y}\right] = \frac{\partial e(u)}{\partial t}\Big|_{u=V(y)}.$$

Furthermore, using the chain rule of differentiation, we obtain

(4.3) $$\frac{\partial e(u)}{\partial t} = \frac{\partial e(u)}{\partial R} \cdot \frac{\partial R(u)}{\partial t}.$$

Hence, (4.2) and (4.3) imply that

(4.4) $$MCF = -\frac{\frac{\partial V(y)}{\partial t} / \frac{\partial V(y)}{\partial y}}{\frac{\partial R(u)}{\partial t}\Big|_{u=V(y)}} = \frac{\frac{\partial e(u)}{\partial t}\Big|_{u=V(y)}}{\frac{\partial R(u)}{\partial t}\Big|_{u=V(y)}} = \frac{\partial e(u)}{\partial R}\Big|_{u=V(y)}.$$

Håkonson (1998), Ballard (1990), and Mayshar (1990) have proposed calculating the marginal cost of public funds by dividing *CV* [6]by the uncompensated change in revenue instead of the compensated one. Jacobs (2018) argues that this approach seems inconsistent because the numerator (*CV*) is a compensated measure, whereas the denominator (change in revenue) is an uncompensated measure.

4.2. The case of discrete choice

To the best of our knowledge there has been no attempt to establish measures of the marginal cost of public funds in the case of discrete labor supply models. We shall now discuss how the results obtained in this paper can be utilized to that end.

Consider the setting of the labor supply model based on the discrete choice framework. Let $T(h, X)$ denote the tax function, as a function of hours of work *h* and the wage rate where *X* is a vector of individual characteristics including non-labor income. Let $\varphi(h \mid X)$ the probability of working *h* hours conditional on *X*. Suppose a policy intervention takes place, consisting of a change in the tax system from *T* to $\tilde{T}$ as the consequence of change in a tax parameter from t to $\tilde{t}$. The corresponding expenditure function $Y(T, \tilde{T}, X)$ is determined by

$$\max_{h\in D}(u(hw - T(h,X) + y, h; X) + \log(\theta(X)g(h)) + \eta(h))$$
$$= \max_{h\in D}(u(hw(X) - \tilde{T}(h,X) + Y(T,\tilde{T},X), h; X) + \log(\theta(X)g(h)) + \eta(h))$$

[6] Hicks (1956) was the first to define the compensating variation, in our context, as $CV = e(u,\tilde{t}) - e(u,t) = e(u,\tilde{t}) - y$, where $\tilde{t}$ and t are ex post and ex ante tax rates.

where we recall that $B(h)$ is the set of jobs with hours of work *h* that are available to the agent and $\{\eta(h), h \in D\}$ are independent draws from the extreme value distribution. Define the marginal aggregate expenditure, $\partial EY(X)/\partial t$, by

(4.5) $$\frac{\partial EY(X)}{\partial t} := \lim_{\tilde{t} \to t} \frac{E(Y(T, \tilde{T}, X) - y)}{\tilde{t} - t}$$

where the expectation operator is taken with respect to both the stochastic terms and *X*. Note that the compensating variation measure $CV(X) = Y(T, \tilde{T}, X) - y$.

Let $R(X)$ denote the (random) revenue in a population with characteristics *X* and let $\partial^{\pm} ER^H(X)/\partial t$ be the left and right marginal compensated expected revenues defined by

(4.6) $$\frac{\partial^{\pm} ER^H(X)}{\partial t} := \sum_{h \in D} \frac{\partial^{\pm} E_X \left(T(h,X)\varphi(h \mid X) \right)}{\partial t}$$

$$= \sum_{h \in D} \left(E_X \left(\frac{\partial T(h,X)}{\partial t} \varphi(h \mid X) \right) + E_X \left(T(h,X) \frac{\partial^{\pm} E\varphi^H(h \mid X)}{\partial t} \right) \right)$$

where E_X is the expectation operator taken with respect to the distribution of *X*.

We propose to define the marginal compensated cost of public funds, $MCF^{\pm}$, by

(4.7) $$MCF^{\pm} := \frac{\partial EY(X)/\partial t}{\partial^{\pm} ER^H(X)/\partial t}$$

where MCF^{-} is the left marginal compensated cost of public funds and MCF^{+} is the right marginal compensated cost of public funds. The marginal cost of public funds given in (4.7) is a way of measuring the aggregate distortionary effect of the tax change within the framework of the discrete choice case based on the random expenditure function. The concept of right and left marginal cost of funds defined in (4.7) are directly measured in units of income, both the numerator and the denominator.[7]

The following result enables us to compute the marginal expected expenditure given in (4.5).

Theorem 4

In the discrete labor supply model

$$\frac{\partial EY(X)}{\partial t} = \sum_{h \in D} E_X \left(\frac{\partial T(h,X)}{\partial t} \varphi(h \mid X) \right).$$

The result in Theorem 4 is similar to the traditional textbook case with continuous hours where hours *h* are now replaced by the probability of working *h* hours $\varphi(h)$. In order to compute the marginal compensated probability $\partial^{\pm}\varphi^H(h \mid X)/\partial t$ we need the following result:

[7] For a recent discussion of applying aggregate money metrics in welfare analysis, see Bosmans et al. (2018).

Theorem 5

In the discrete labor supply model the marginal compensated effects of the labor supply choice probabilities with respect to the tax parameter t is given by

$$\frac{\partial^{+}\varphi^{H}(h\,|\,X)}{\partial t}=\varphi(h\,|\,X)\frac{\partial v(h;X)}{\partial y}\sum_{x\in D}\varphi(x\,|\,X)(\xi(x,X)-\xi(h,X))_{+}$$

$$-\varphi(h\,|\,X)\sum_{x\in D}\varphi(x\,|\,X)\frac{\partial v(x;X)}{\partial y}(\xi(h,X)-\xi(x,X))_{+}-\varphi(h\,|\,X)\varphi(0\,|\,X)\xi(h,X)\frac{\partial v(h;X)}{\partial y}$$

and

$$\frac{\partial^{-}\varphi^{H}(h\,|\,X)}{\partial t}=\varphi(h\,|\,X)\frac{\partial v(h;X)}{\partial y}\sum_{x\in D}\varphi(x\,|\,X)(\xi(x,X)-\xi(h,X))_{+}$$

$$-\varphi(h\,|\,X)\sum_{x\in D}\varphi(x\,|\,X)\frac{\partial v(x;X)}{\partial y}(\xi(h,X)-\xi(x,X))_{+}-\varphi(h\,|\,X)\varphi(0\,|\,X)\xi(h,X)\frac{\partial v(0;X)}{\partial y}$$

where $v(h;X)=u(hw(X)-T(h,X)+y,h;X)+\log(\theta(X)g(h))$ *and* $\xi(h,X)=\partial T(h,X)/\partial t.$

The result of Theorem 5 follows from Theorem 2 with $\gamma(h)=-\xi(h,X).$ From Theorem 5 we observe that

$$\frac{\partial^{+}\varphi^{H}(h\,|\,X)}{\partial t}-\frac{\partial^{-}\varphi^{H}(h\,|\,X)}{\partial t}=\varphi(h\,|\,X)\varphi(0\,|\,X)\xi(h,X)\left(\frac{\partial v(0;X)}{\partial y}-\frac{\partial v(h;X)}{\partial y}\right),$$

which can attain both positive and negative values.

In practice, changes in the tax system are rarely approximately infinitesimal. It is therefore of interest to apply a measure that can be used when it is relevant to evaluate an alternative tax system, where it is understood that the slope of several tax segments and hence their boundaries change. In such cases the formula in (4.7) cannot be used. Instead one should apply

$$\text{(4.8)}\qquad MCF=\frac{ECV(X)}{\Delta ER^{H}(X)}$$

where we recall that $CV(X)$ captures the actual change in expected expenditure implied by the change to the alternative tax system and $\Delta ER^{H}(X)$ is the corresponding change in the compensated expected revenue. In order to calculate $\Delta ER^{H}(X)$ one can apply the results of Theorem 1. To this end let $\tilde{f}(x,z)=x+z-\tilde{T}(x,z),$ $v_j(h;y,X)=u(f(hw(X),y),h,X)+\log(\theta(X)g(h)),$ $\tilde{v}_j(h;y,X)=u(\tilde{f}(hw_j,y),h,X)+\log(\theta(X)g(h))$ and let y_{jh} be determined by $v(h;y,X)=v(h;y_h(X),X)$ and $v(0;y,X)=\tilde{v}(0;y_0(X),X).$ It then follows from Theorem 1 that

$$\text{(4.9)}\qquad E(Y(X)\,|\,X)=\sum_{h\in D\setminus\{0\}}\int_{0}^{y_h(X)}\frac{\exp(v(h;z,X))dz}{M(z,X)}+\int_{0}^{y_0(X)}\frac{\exp(v(0;z,X))dz}{M(z,X)}$$

where

$$M(z,X)=\sum_{x\in D}\exp(\max\big(v(x;y,X),\tilde{v}(x;z,X)\big)).$$

Let $Q^H(x,h\,|\,X)$ be the compensated probability of choosing hours *x* ex ante and hours *h* ex post, conditional on *X*. These probabilities are given in Theorem 1. Moreover, let $P^H(h\,|\,X)$ be the ex post compensated probability of working *h* hours. Evidently,

$$P^H(h\,|\,X)=\sum_{x\in D}Q^H(x,h\,|\,X).$$

Hence, it follows that

(4.10) $$\Delta ER^H(X)=\sum_{h\in D}E_X\left(\tilde{T}(h,X)P^H(h\,|\,X)-T(h,X)\varphi(h\,|\,X)\right).$$

The formulas given in (4.9) and (4.10) can easily be extended to the multisectoral case by using Lemma B.1 in Appendix B.

5. Numerical results of compensated and uncompensated wage elasticities

To illustrate how the compensated and uncompensated elasticities differ across individual characteristics, we report labor supply wage and income elasticities based on the empirical two-sector discrete labor supply model estimated by Dagsvik and Strøm (2006). The two sectors are public and private. The model is estimated on Norwegian data from 1994 and 1995. We consider only married women. Data, tax functions, and estimates of the deterministic part of the utility function and of θ_j and $g_j(h)$ are given in Appendix C[8]: see Dagsvik and Strøm (2006) for more details. All values below refer to 1994. Since then the wage level has increased substantially. The average wage per hour for married women in 1994 was around NOK 90–110, while the average in 2018 was around NOK 230–260. Wages in the private sector were the highest ones. As of January 2019, 1 USD is equivalent to approximately NOK 8.5.

The deterministic part of the utility function used in the estimation is given by

(5.1) $$u(f(hw_j,y)),h,X)=\frac{(10^{-4}(f(hw_j,y)-C_0))^{\alpha_1}-1}{\alpha_1}\left[\alpha_2+\alpha_9\frac{(1-h/3640)^{\alpha_3}-1}{\alpha_3}\right]+Xb\frac{(1-h/3640)^{\alpha_3}-1}{\alpha_3}.$$

[8] Appendices C-J are given in the supplementary section, available at http://folk.uio.no/steinast/supplements/Supplement%20Compensated%20Marginal%20Effects%20and%20the%20Slutsky%20Equation%20in%20Discrete%20Labor%20Supply%20Models.pdf

where C_0 is the subsistence consumption level, and X is a vector of observed characteristics: number of children 0–6 and 7–17, log age, and log age squared. The term $1-h/3460$ is a normalized expression for leisure after taking time for rest and sleep into consideration. The estimates imply that the utility function is quasi-concave, with a significant negative interaction term for leisure and consumption.

We have computed wage and income elasticities for selected groups of married women who differ with respect to given wage levels (low, medium high, and very high), non-labor income (low, medium high, and very high), number of children (none or two), and age (30 and 40). In total we have 36 different cases. A summary of the different cases is given in Table 5.1 below.

Table 5.1. The selected households

1–3	woman aged 30,	wage 70,	no children,	not labor income 70,000, 100,000, 200,000	respectively
4–6	woman aged 30,	wage 200,	no children,	not labor income 70,000, 100,000, 200,000	respectively
7–9	woman aged 30,	wage 300,	no children,	not labor income 70,000, 100,000, 200,000	respectively
10–12	woman aged 30,	wage 70,	two children,	not labor income 70,000, 100,000, 200,000	respectively
13–15	woman aged 30,	wage 200,	two children,	not labor income 70,000, 100,000, 200,000	respectively
16–18	woman aged 30,	wage 300,	two children,	not labor income 70,000, 100,000, 200,000	respectively
19–21	woman aged 40,	wage 70,	no children,	not labor income 70,000, 100,000, 200,000	respectively
22–24	woman aged 40,	wage 200,	no children,	not labor income 70,000, 100,000, 200,000	respectively
25–27	woman aged 40,	wage 300,	no children,	not labor income 70,000, 100,000, 200,000	respectively
28–30	woman aged 40,	wage 70,	two children,	not labor income 70,000, 100,000, 200,000	respectively
31–33	woman aged 40,	wage 200,	two children,	not labor income 70,000, 100,000, 200,000	respectively
34–36	woman aged 40,	wage 300,	two children,	not labor income 70,000, 100,000, 200,000	respectively

The non-labor income is the sum of capital income after tax, child allowances, and the after-tax income of the husband. We account for the fact that when the woman has no income or a very low income, she is taxed together with her husband. The tax functions are given in Appendix C. The probabilities of participation at all and in the two sectors, and expected working hours for these 36 cases are given in Appendix D.

In Section 5.1 we report the uncompensated and the compensated elasticities when the wage rate in the public and the private sectors respectively is increased (right elasticities). Next, in Section 5.2 we compare the right and left elasticities. In Section 5.3 we show the elasticities when there is an overall wage change (both sectors and all alternatives), again for the same 36 women, while in Section 5.4 we report the elasticities when there is an overall wage change using the whole sample from 1994 to 1995. In this last case, the elasticities are aggregate in the sense that they are based on the aggregate expected hours of work where the aggregation takes place over individual characteristics, non-labor incomes, and wage rates. The individual wage rates are represented by a standard-type wage equation with normally distributed error terms that are assumed to be independent of the taste shifters in the utility function. The empirical part of the paper ends in Section 5.5, where we give an estimate of the marginal cost of public funds based on the sample from 1994. Numerical values for all elasticities are given in Appendix E–J.

5.1. Sector-specific wage increases and elasticities for selected women

Table 5.2 provides an example of the marginal *right* compensated and uncompensated wage elasticities. The effects in all 36 cases when the *public-sector wage* only is increased is given in Appendix F. In Appendix G we give the results when the private-sector wage is increased. Table 5.2 shows that the compensated and uncompensated cross elasticities are equal. Also the direct elasticities, compensated and uncompensated, of the probability of working in a sector are equal. The compensated direct elasticities of expected hours are always higher than the corresponding uncompensated elasticities. The total elasticities are the net of direct and cross effects.

That the cross elasticity of *conditional expected hours*, conditional on working in a specific sector, is equal to zero is seen as follows. The elasticity of the conditional expectation of hours of work is given by

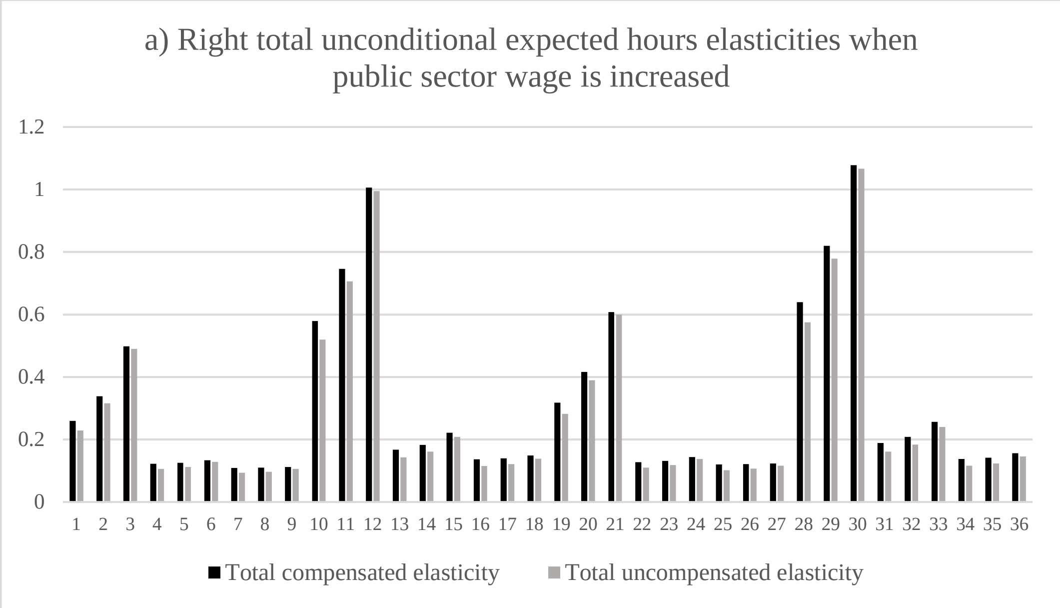
$$\partial \log\left(\frac{\sum_{x\in D} x\varphi_k(x)}{\sum_{x\in D}\varphi_k(x)}\right) / \partial \log w_j = \frac{w_j}{\sum_{x\in D} x\varphi_k(x)}\cdot\sum_{x\in D} x\frac{\partial\varphi_k(x)}{\partial w_j} - \frac{w_j}{\sum_{x\in D}\varphi_k(x)}\cdot\sum_{x\in D}\frac{\partial\varphi_k(x)}{\partial w_j}$$

$$= -\frac{w_j}{\sum_{x\in D} x\varphi_k(x)}\cdot\sum_{x\in D} x\varphi_k(x)\sum_{x\in D}\varphi_j(x)\frac{\partial v_j(x)}{\partial w_j} + \frac{w_j}{\sum_{x\in D}\varphi_k(x)}\cdot\sum_{x\in D}\varphi_k(x)\sum_{x\in D}\varphi_j(x)\frac{\partial v_j(x)}{\partial w_j} = 0.$$

In Figure 5.1 we show the elasticities of the *unconditional expectation of total hours* with respect to an increase in the public wage (graph a) and the private wage (graph b) for the selected 36 women described in Table 5.1. “Total” means that net elasticity takes into account cross effects. The compensated elasticities exceed the uncompensated elasticities, but the difference is almost negligible. The elasticities are small, around 0.2, with the exception of women with very low wages.

Table 5.2. Right uncompensated and compensated wage elasticities. Married woman aged 30, no children, hourly wage NOK 70

Income		Probability of working			Conditional hours			Unconditional hours		
		All	Public (direct)	Private (cross)	Total	Public (direct)	Private (cross)	Total	Public (direct)	Private (cross)
70,000	Uncompensated	0.008	1.442	-5.753	0.221	0.236	0.000	0.229	1.679	-5.753
	Compensated	0.008	1.442	-5.753	0.252	0.275	0.000	0.260	1.717	-5.753
100,000	Uncompensated	0.035	1.341	-5.094	0.281	0.303	0.000	0.316	1.643	-5.094
	Compensated	0.035	1.341	-5.094	0.303	0.330	0.000	0.338	1.671	-5.094
200,000	Uncompensated	0.131	1.184	-3.839	0.359	0.396	0.000	0.490	1.580	-3.839
	Compensated	0.131	1.184	-3.839	0.367	0.405	0.000	0.498	1.589	-3.839



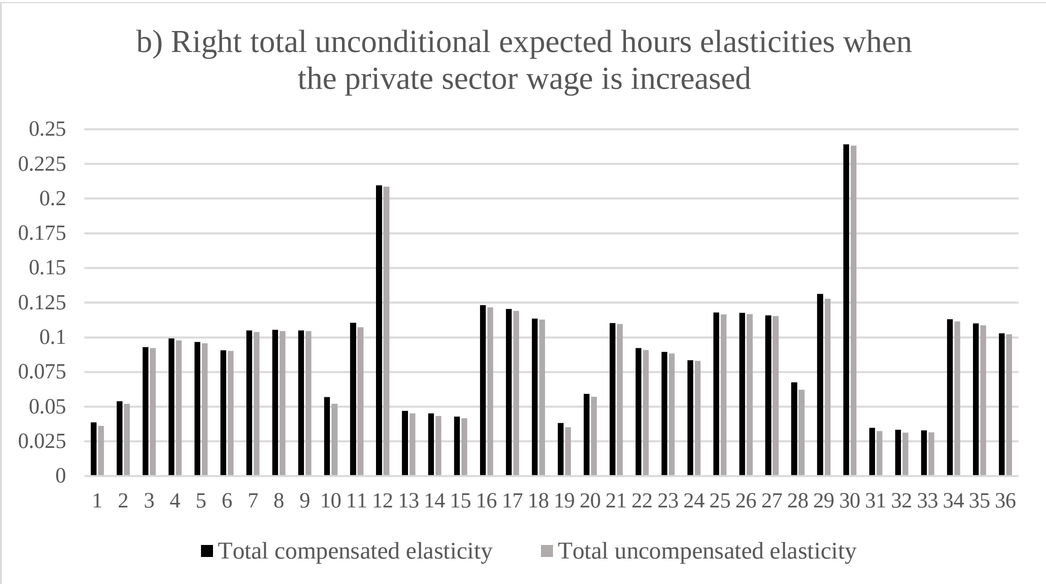
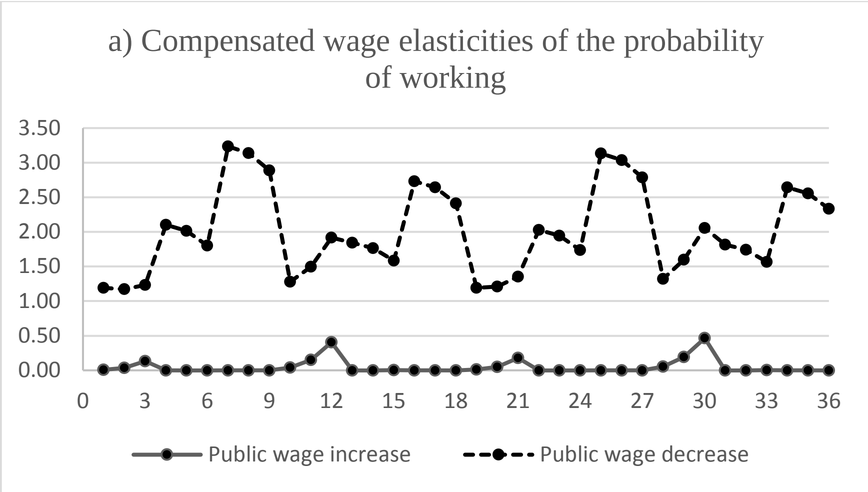


Figure 5.1. Total unconditional expected hours elasticities related to one sector wage increase

5.2. Sector-specific wage increases versus wage decreases and elasticities

In Appendix H (public-sector wage decrease) and Appendix I (private-sector wage decrease) we report left wage elasticities for all 36 types of women. Figure 5.2 (a, b) give the total wage elasticities of the probability of working (extensive margin). “Total” means that cross effects are accounted for. Figure 5.3 (a, b) give the elasticities of the conditional mean hours (intensive margin) and, finally, Figure 5.4 (a, b) give the elasticities of the unconditional expected hours of work.

Note that at the extensive margin the left total wage elasticities are much higher than the right total wage elasticities. At the intensive margin the differences between left and right total wage elasticities are almost negligible for the public-sector wage and somewhat larger, but not much, for the private-sector wage.



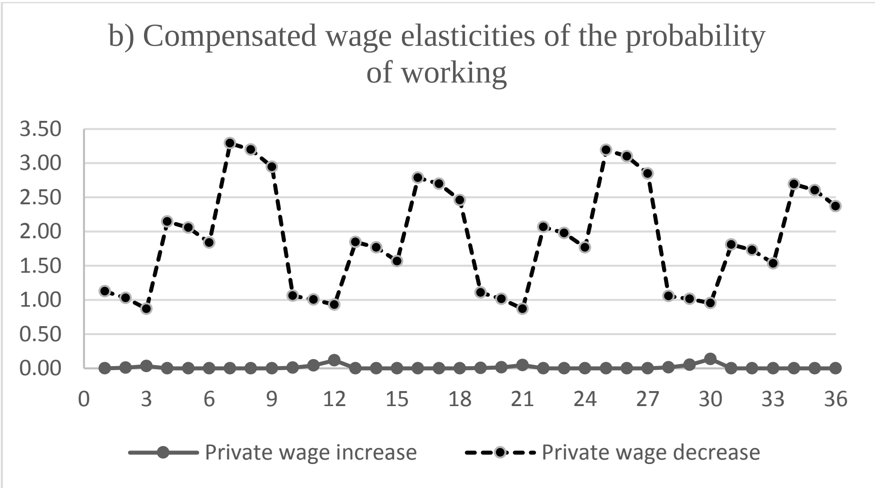
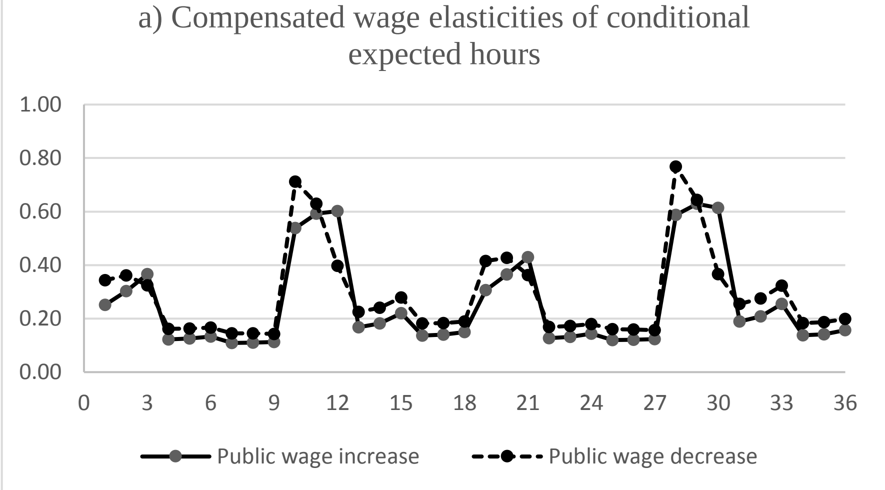


Figure 5.2. Total compensated wage elasticities of the probability of working



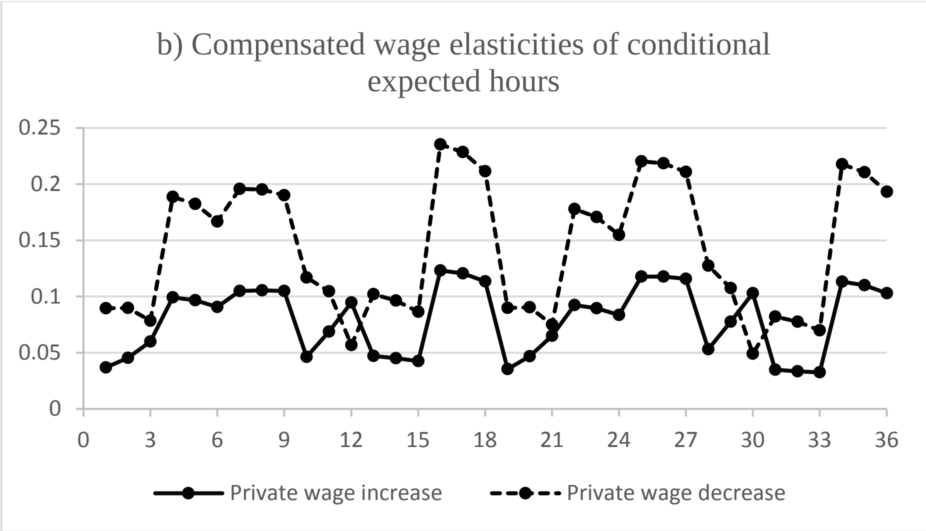
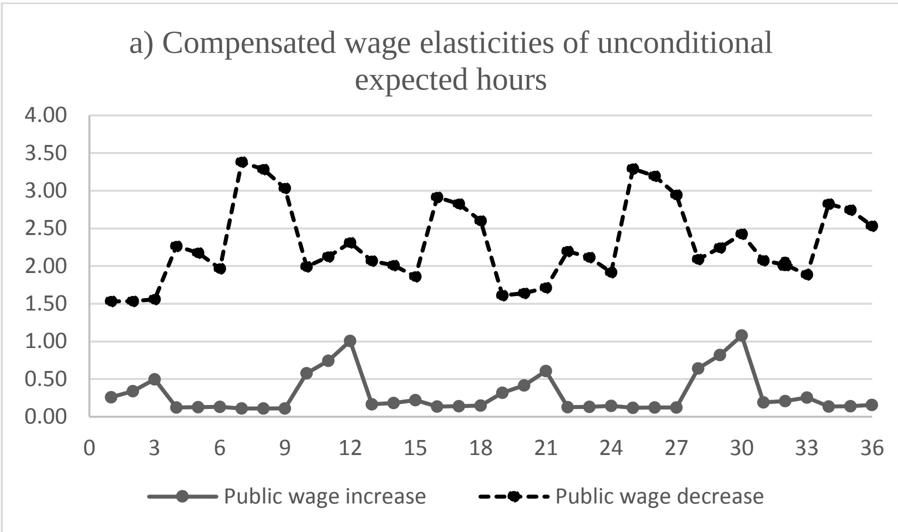


Figure 5.3. Total compensated wage elasticities of conditional expected hours



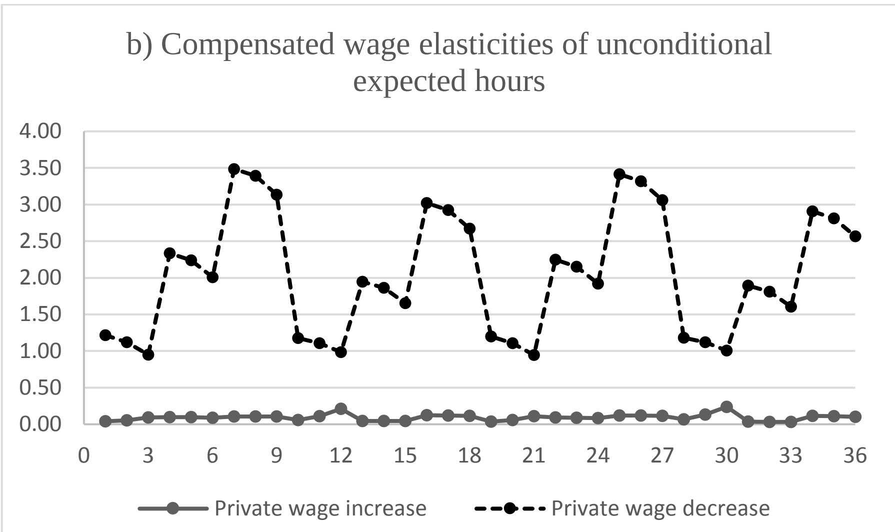


Figure 5.4. Total compensated wage elasticities of unconditional expected hours

5.3. An overall wage increase and elasticities for selected women

To simplify the exposition, we consider the case where the wage rates are the same in both sectors, which imply that $\min(a_j\gamma_j(h), a_k\gamma_k(x)) = \min(\gamma(h), \gamma(x))$ and $\max(a_j\gamma_j(h), a_k\gamma_k(x)) = \max(\gamma(h), \gamma(x))$; see Corollaries A3 to A6 in Appendix A. The difference between the compensated and the uncompensated wage elasticities can be calculated using the formulas given in Section 3 and in Appendix A. In Table 5.3 we report the results for one case. The other 35 cases are given in Appendix E. As shown in Section 3 and Appendix A, the uncompensated and compensated wage elasticities of the probability of working at all are equal. The compensated wage elasticities of unconditional hours are larger than the uncompensated ones. The wage elasticities of participation at all are low, which is because the participation probabilities are high (close to 1).

The higher the non-labor income, the lower is the probability of participation and the higher is the wage elasticities. Given the woman's wage rate level, we observe that this is also the case for the wage elasticities of mean hours of work. We observe that the elasticities at the intensive margin are clearly higher than at the extensive margin. We also note that an overall wage increase shifts labor from the private to the public sector.

Table 5.3. An overall wage increase: right compensated and uncompensated wage elasticities. Married woman aged 30, no children, hourly wage NOK 70

Non-labor income	Direct Elasticity	Probability of working			Conditional mean hours			Unconditional mean hours		
		All	Public	Private	All	Public	Private	All	Public	Private
70,000	Uncompensated	0.010	0.047	-0.140	0.255	0.236	0.330	0.265	0.283	0.190
	Compensated	0.010	0.054	-0.170	0.307	0.284	0.397	0.317	0.339	0.226
100,000	Uncompensated	0.044	0.090	-0.139	0.325	0.303	0.407	0.368	0.393	0.269
	Compensated	0.044	0.096	-0.162	0.362	0.337	0.454	0.406	0.433	0.292
200,000	Uncompensated	0.164	0.219	-0.042	0.419	0.396	0.498	0.583	0.615	0.456
	Compensated	0.164	0.221	-0.050	0.431	0.408	0.512	0.595	0.629	0.462

In Figure 5.5 we show the aggregate elasticities of unconditional expected hours elasticities for the 36 different cases of women, aggregated across sectors. The wage elasticities are relatively small, with 12 exceptions. All 12 have low hourly wages (NOK 70, which is somewhat lower than the average hourly wage in 1994). Among them the wage elasticities are higher if they have children and/or if the non-wage income is high. These results clearly demonstrate that heterogeneity matters.

The results shown in Figure 5.5 are in line with what Attanasio et al. (2018) have found. The model they use is, however, quite different from our model.

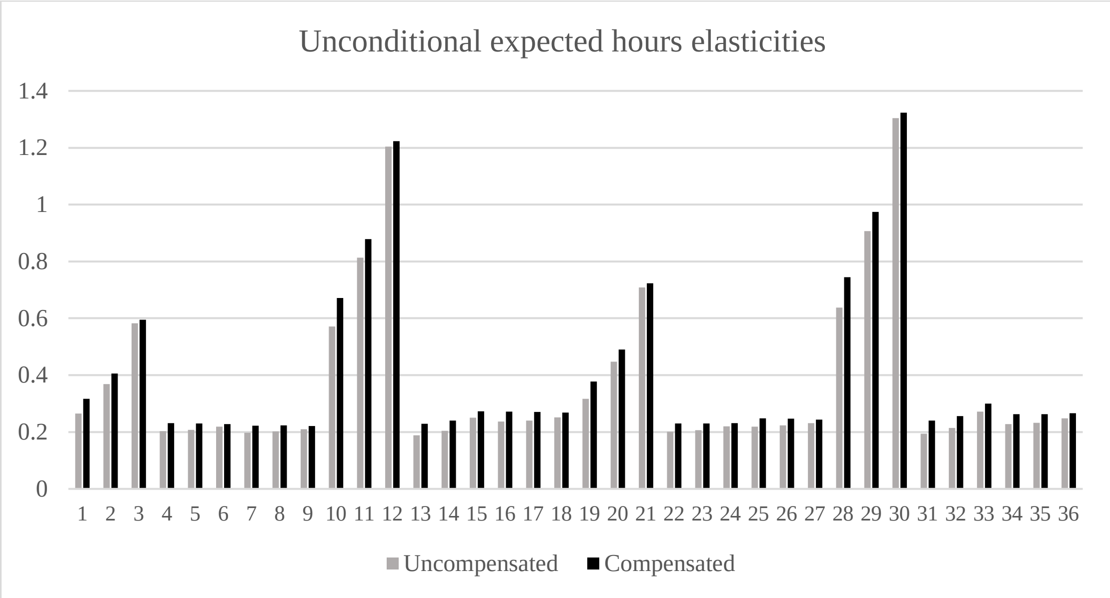


Figure 5.5. Right compensated and uncompensated wage elasticities of total unconditional expected hours for an overall wage increase

5.4. Overall wage change and aggregate elasticities using sample data

In this section we report aggregate wage elasticities based on the sample used to estimate the model. In the empirical model, the wage rates are represented by a wage equation which is integrated out. Table 5.4 displays the right compensated and uncompensated expected elasticities of an *overall* increase in the wage rates on the probability of working at all, working in the public and the private sector respectively, on mean hours conditional on participation, and on the unconditional mean hours.

From Table 5.4 we note that the uncompensated and right compensated elasticities of working at all are equal, whereas the right compensated elasticity of total hours of an overall wage change, conditional on working and hence also in the unconditional case, exceeds the corresponding uncompensated elasticity. However, the differences between the compensated and uncompensated elasticities are not large. As mentioned in Section 3 these differences are the respective income effects.

We note that the elasticity of mean hours is clearly larger for the private sector than for the public sector. The reason why labor supply related to working in the private sector is more elastic than working in the public sector is that there are fewer constraints on offered hours in the private sector (captured by the opportunity distributions $\theta_j g_j(h)$, j = 1,2, which represent the latent choice set). The unconditional hours wage elasticities are also given in Figure 5.6.

Table 5.4. Right compensated and uncompensated wage elasticities of an overall wage increase using sample values. Married women, Norway 1994

	Probability of working			Conditional hours, given working			Unconditional hours		
	Total	Public	Private	Total	Public	Private	Total	Public	Private
Uncompensated elasticity	0.3149	0.1071	0.5186	0.3724	0.3896	0.3489	0.6873	0.4967	0.8675
Compensated elasticity	0.3149	0.0740	0.5510	0.4012	0.4130	0.3817	0.7161	0.4870	0.9327

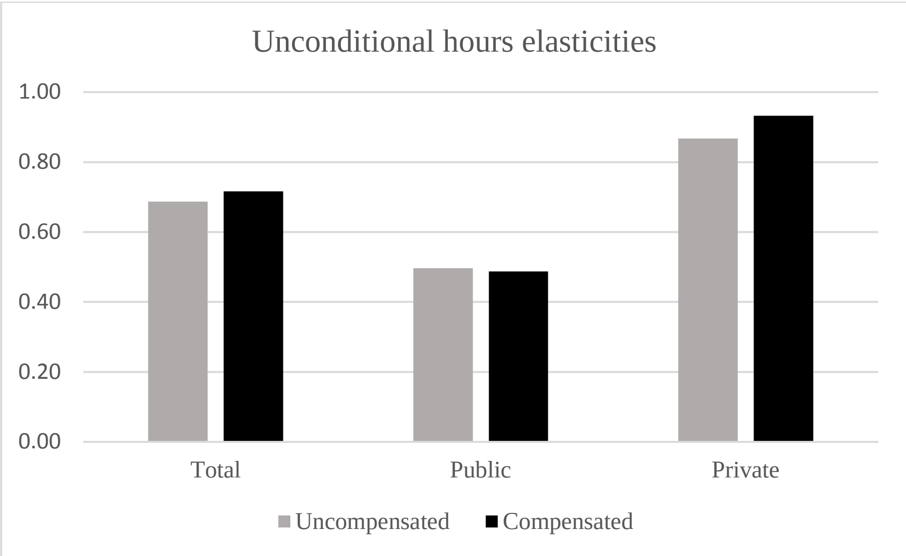


Figure 5.6. **Elasticities of the unconditional expectation of hours with respect to an overall wage increase**

5.5. The marginal cost of funds from an increase in the minimum deduction in the tax function

In this Section we apply the results of Section 4 to compute the marginal cost of funds that follows from an increase in the minimum deduction in the tax system. Like many countries, Norway has a progressive tax system with regard to the taxation of labor income, with stepwise linear parts. Let the marginal tax rates be denoted t_k where the subscript *k* refers to tax segment *k*. Let I_0 denote the level of the minimum deduction. The actual structure of the tax system in 1994 is given in Appendix C.

In order to make a change in the tax rates of a stepwise linear tax structure it is necessary, whenever a tax rate is changed, to change the tax segment boundaries too. This implies that to assess the impact of a tax rate change one has to make a discrete change in the tax structure. This is not required when the minimum deduction level I_0 is changed. The uncompensated and compensated marginal effects of a change in the minimum deduction I_0 (below this level taxes are zero) are zero for $w_j(X)h \le I_0$ and

$$\frac{\partial \varphi_j(h|X)}{\partial I_0} = t_1 \varphi_j(h|X)\left[\frac{\partial v_j(h,X)}{\partial y} - \sum_r \sum_{xw_r(X)>I_0} \varphi_r(x|X)\frac{\partial v_r(x,X)}{\partial y}\right] \text{ for } w_j(X)h \ge I_0.$$

In this case the right and left compensated effects are equal and the Slutsky equations are given by

$$\frac{\partial \varphi_j(h|X)}{\partial I_0} = \frac{\partial \varphi_j^H(h|X)}{\partial I_0}$$

for $w_j(X)h < I_0$ and the Slutsky equation is

$$\frac{\partial \varphi_j(h|X)}{\partial I_o} = \frac{\partial \varphi_j^H(h|X)}{\partial I_0} - t_1 \varphi_j(h|X)\sum_k \sum_{w_k(X)x>I_0} \varphi_k(x|X)\left(\frac{\partial v_k(x,X)}{\partial y} - \frac{\partial v_j(h,X)}{\partial y}\right)$$

for $w_j(X)h \ge I_0$.

To compute the marginal costs of funds we apply the model given in Dagsvik and Strøm (2006) estimated on a sample of Norwegian married women. In the Norwegian tax system the marginal tax at the lowest tax bracket is equal to $t_1 = 0.302$. Hence, for $w_j h > I_o$

$$\frac{\partial T_j(h)}{\partial I_0} = -t_1 = -0.302.$$

By a straight forward extension of Theorem 4 it follows that, similarly to (4.8), MCF can be computed using the following formula:

$$MCF = \frac{\sum_{i\in S}\sum_j \sum_{h\in D} \frac{\partial T_j(h,X_i)}{\partial I_0}\varphi_j(h|X_i)}{\sum_{i\in S}\sum_j \sum_{h\in D}\left(\frac{\partial T_j(h,X_i)}{\partial I_0}\varphi_j(h|X_i) + T_j(h,X_i)\frac{\partial^+ \varphi_j^H(h|X_i)}{\partial I_0}\right)}$$

where S denotes the sample of married women, $\varphi_j(0) = 0$ when $j > 0$, $\varphi_0(h) = 0$ when $h > 0$ and $\varphi_0(0) = \varphi(0)$. This measure of the marginal cost of public funds assumes implicitly that taxes are optimized and distributional concerns are absent. This means that we assume the government has done its best to optimize taxes and redistribute income. We have no reasons to overturn the judgment of the politicians (and the Ministry of Finance). Note that both the numerator and the denominator are expressed in money terms and are summed over individuals. We also assume that lump-sum taxes are not an alternative.

Our estimate of the marginal cost of public funds is 1.15, which is in the lower part of the range that others have found, also in Norway, as referred to above. It should be noted that since we include only married women in our sample, our estimate is most likely to be higher than if single women and men (single and married) had been included. Numerous studies have demonstrated that married women respond more strongly to changes in economic incentives than single women and men, married or not.

6. Conclusions

Using the results of Dagsvik and Karlström (2005) as a starting point, we have derived analytic formulas for marginal compensated wage effects in discrete labor supply models. In particular, we have established Slutsky-type relations for these kinds of models. These relations differ in important ways from the traditional Slutsky equations. In particular, the left marginal compensated wage effects differ in general from the corresponding right marginal compensated effects. In the selected numerical examples we have used, it turns out that this difference is not large at the intensive margin, but rather sizeable at the extensive margin. In the selected numerical examples, the compensated labor supply elasticities are larger than the uncompensated ones, but the difference is small, indicating minor income effects. The labor supply elasticities, compensated as well as uncompensated, are greatest for those married women with the lowest wage, those with many children, and those with the highest non-labor income. Apparently heterogeneity matters with respect to responses to wage and tax rate changes.

Finally, we have discussed how marginal cost of funds can be calculated based on discrete labor supply models. Our estimate of the marginal cost of public funds is 1.15, which is lower than the one used in current cost-benefit analysis in Norway.

Funding: This study was funded by the Ragnar Frisch Centre of Economic Research, Oslo, Norway, and by Small Research Grants from the Department of Economics, University of Oslo.

Conflict of Interest: John K. Dagsvik has received research grants from the Ragnar Frisch Centre of Economic Research, Oslo, Norway. Marilena Locatelli has received research grants from the Ragnar Frisch Centre of Economic Research, Oslo, Norway.Steinar Strøm has received research grants from the Ragnar Frisch Centre of Economic Research, Oslo, Norway. The authors declare that they have no conflict of interest.

References

Attanasio, O., P. Levell, H. Low and V. Sanchez-Marcos (2018): Aggregating Elasticities: Intensive and Extensive Margins of Women's Labor Supply. *Econometrica*, **86**, 2049-2082.

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

Aaberge, R., U. Colombino and S. Strøm (1999): Labor Supply in Italy. An Empirical Analysis of Joint Household Decisions, with Taxes and Quantity Constraints. *Journal of Applied Econometrics*, **14**, 403-422.

Atkinson, A. B. and N. H. Stern (1974): Pigou, Taxation and Public Goods, *Review of Economic Studies*, **41**, 119-128.

Ballard, C.L. (1990): Marginal Welfare Cost Calculation: Differential Analysis vs Balanced-Budget Analysis. *Journal of Public Economics*, **41**, 263-276.

Bosmans, K., K. Decancq and E. Ooghe (2018): Who's Afraid of Aggregating Money Metrics? *Theoretical Economics*, **13**, 467-484.

Browning, E. K. (1976): The Marginal Cost of Public Funds, *Journal of Political Economy*, **84**, 283-298.

Browning, E. K. (1987): On the Marginal Welfare Cost of Taxation. *American Economic Review*, **77**, 11-23.

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

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

Dagsvik, J. K. (2000): Aggregation in Matching Markets. *International Economic Review*, **41**, 27–57.

Dagsvik, J. K. (2018): Discrete Choice, Compensated Effects and the Slutsky Equation. Mimeo, to appear.

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**, 57-76.

Dagsvik, J. K., O. Ljones, S. Strøm and T. Wennemo (1988): The Impact on Labor Supply of a Shorter Workday. A Microeconometric Discrete/Continuous Choice Approach. In R. A. Hart (ed.): *Employment, Unemployment and Labor Utilization*. Unwin Hyman, London.

Dagsvik, J. K. and S. Strøm (1992): Labor Supply with Non-convex Budget Sets, Hours Restriction and Non-pecuniary Job Attributes. Discussion Papers no. 76, Statistics Norway.

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

Dagsvik, J. K. 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. and Z. Jia (2018): Aggregate Behavior in Matching Markets with Flexible Contracts and Non-transferable Representations of Preferences. Discussion Papers no 875, Research Department, Statistics Norway.

Dagsvik, J. K., M. Locatelli and S. Strøm (2009): Tax Reform, Sector Specific Labor Supply and Welfare Effects. *Scandinavian Journal of Economics,* **111**, 265-287.

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.

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

Harberger, A.C. (1964): The Measurement of Waste. *American Economic Review*, **54**, 58-76.

Hicks, J. R. (1936): *Value and Capital*. Oxford University Press.

Hicks, J.R. (1956): *A Revision of Demand Theory*, Oxford University Press.

Håkonson, L. (1998): An Investigation into Alternative Representation of the Marginal Cost of Public Funds. *International Tax Public Finance*, **5**, 329-343.

Jacobs, B. (2018): The Marginal Cost of Public Funds. *International Tax Public Finance*, **25**, 883-912.

Kleven, H.J. and C.T. Kreiner (2006): The Marginal Cost of Public Funds: Hours of Work versus Labor Force Participation. *Journal of Public Economics*, **90**, 1955-1973.

Mayshar, J. (1990): On Measures of Excess Burden and Their Application. *Journal of Public Economics*, **43**, 263-289.

Menzel, K. (2015): Large Matching Markets as Two-sided Demand Systems. *Econometrica*, **83**, 897-941.

Pigou, A. C. (1947): *A Study in Public Finance*. Macmillan, London.

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

Slutsky, E. E. (1915): Sulla Teoria del Bilancio del Consumatore. *Giornale degli Economisti*, **51**, 1–26.

Small, K. A. and H. S. Rosen (1981): Applied Welfare Analysis in Discrete Choice models. *Econometrica*, **49**, 105-130.

Statistics Norway (1994), Survey of Income and Wealth. www.ssb.no/emner/05/01/.

Statistics Norway (1995), Level of Living Conditions. www.ssb.no/emner/00/02/.

Stiglitz, J. and P. S. Dasgupta (1971): Differential Taxation, Public Goods and Economic Efficiency. *Review of Economic Studies*, **38**, 151-174.

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

Varian, H. (1992): *Microeconomic Analysis* (3rd ed.). W. W. Norton, New York.

Appendix A. Further detailed formulas for various marginal effects

Theorem A1

Under the assumptions of the multisectoral discrete labor supply model we have

$$\frac{\partial^+\varphi_j^H(h)}{\partial w_j}=\frac{\partial v_j(h)}{\partial y}\varphi_j(h)\sum_{x\in D\setminus\{0\}}\varphi_j(x)(\gamma_j(h)-\gamma_j(x))_+-\varphi_j(h)\sum_{x\in D\setminus\{0\}}\varphi_j(x)\frac{\partial v_j(x)}{\partial y}(\gamma_j(x)-\gamma_j(h))_+$$

$$+\varphi_j(h)(1-\varphi_j)\frac{\partial v_j(h)}{\partial y}\gamma_j(h),$$

$$\frac{\partial^-\varphi_j^H(h)}{\partial w_j}=\varphi_j(h)\sum_{x\in D\setminus\{0\}}\varphi_j(x)\frac{\partial v_j(x)}{\partial y}(\gamma_j(h)-\gamma_j(x))_+-\varphi_j(h)\frac{\partial v_j(h)}{\partial y}\sum_{x\in D\setminus\{0\}}\varphi_j(x)(\gamma_j(x)-\gamma_j(h))_+$$

$$+\varphi_j(h)\gamma_j(h)\sum_{r\neq j,0}\sum_{x\in D\setminus\{0\}}\frac{\partial v_r(x)}{\partial y}\varphi_r(x)+\varphi_j(h)\gamma_j(h)\varphi(0)\frac{\partial v(0)}{\partial y},$$

$$\frac{\partial^+\varphi_j^H(h)}{\partial w_k}=\frac{\partial\varphi_j(h)}{\partial w_k}\quad and\quad \frac{\partial^-\varphi_j^H(h)}{\partial w_k}=-\varphi_j(h)\frac{\partial v_j(h)}{\partial y}\sum_{x\in D\setminus\{0\}}\varphi_k(x)\gamma_k(x)$$

for $k\neq j,\ j,k>0,$ *and*

The proof of Theorem A1 is given in Appendix B.

Corollary A1

Let $\tilde h_j$ *and* $\tilde h_j'$ *be independent draws from the labor supply probability mass function* $\varphi(h)$. *Then under the assumptions of Theorem A1 with* $j>0$

$$\frac{\partial^+E\tilde h_j^H}{\partial w_j}-\frac{\partial E\tilde h_j}{\partial w_j}=Cov\left(\tilde h_j,-\frac{\partial v_j(\tilde h_j)}{\partial y}\min\left(\gamma_j(\tilde h_j),\gamma_j(\tilde h_j')\right)\right)-Cov\left(\tilde h_j',-\frac{\partial v_j(\tilde h_j)}{\partial y}\min\left(\gamma_j(\tilde h_j),\gamma_j(\tilde h_j')\right)\right),$$

$$\frac{\partial^-E\tilde h_j^H}{\partial w_j}-\frac{\partial E\tilde h_j}{\partial w_j}=Cov\left(\tilde h_j,-\frac{\partial v_j(\tilde h_j)}{\partial y}\max\left(\gamma_j(\tilde h_j),\gamma_j(\tilde h_j')\right)\right)-Cov\left(\tilde h_j',-\frac{\partial v_j(\tilde h_j)}{\partial y}\max\left(\gamma_j(\tilde h_j),\gamma_j(\tilde h_j')\right)\right)$$

$$+\sum_{r\neq j}E\left(\frac{\partial v_r(\tilde h_r)}{\partial y}\tilde h_j\gamma_j(\tilde h_j)\right)-(1-\varphi_j)E\left(\frac{\partial v_j(\tilde h_j)}{\partial y}\tilde h_j\gamma_j(\tilde h_j)\right)$$

and

$$\frac{\partial^-E\tilde h_j^H}{\partial w_k}-\frac{\partial E\tilde h_j}{\partial w_k}=E\tilde h_jE\left(\frac{\gamma_k(\tilde h_k)\partial v_k(\tilde h_k)}{\partial y}\right)-E\gamma_k(\tilde h_k)E\left(\frac{\tilde h_j\partial v_j(\tilde h_j)}{\partial y}\right).$$

The proof of Corollary A1 is given in Appendix B.

Corollary A2

Under the Assumptions of Theorem 3 we have that

$$\frac{\partial^+ E\tilde{h}_j^H}{\partial w_j} > \frac{\partial E\tilde{h}_j}{\partial w_j}$$

The proof of Corollary A2 is given in Appendix B. In contrast to the single sector case we have not been able to prove that in general

$$\frac{\partial^- E\tilde{h}_j^H}{\partial w_j} > \frac{\partial E\tilde{h}_j}{\partial w_j}.$$

However, this inequality holds for our empirical model, cf. section 5.2.

The formulas above concern the case where only sector specific wages are altered. It is, however, also of interest to consider the case with an overall wage increase. Specifically, assume now that $w_j = a_j w$ where w is a wage component that is common to both sectors. Then we have that for $j > 0$.

Corollary A3

Under the assumptions of Theorem A1 and with $w_j = a_j w$ where w is a common wage component, we have for h > 0,

$$\frac{\partial^+ \varphi_j^H(h)}{\partial w} = \varphi_j(h)\frac{\partial v_j(h)}{\partial y}\sum_k \sum_{x\in D\backslash\{0\}} \varphi_k(x)(a_j\gamma_j(h) - a_k\gamma_k(x))_+$$

$$-\varphi_j(h)\sum_k \sum_{x\in D\backslash\{0\}} \varphi_k(x)\frac{\partial v_k(x)}{\partial y}(a_k\gamma_k(x) - a_j\gamma_j(h))_+ + \varphi_j(h)\varphi(0)a_j\gamma_j(h)\frac{\partial v_j(h)}{\partial y}$$

and

$$\frac{\partial^- \varphi_j^H(h)}{\partial w} = \varphi_j(h)\sum_k \sum_{x\in D\backslash\{0\}} \frac{\partial v_k(x)}{\partial y}\varphi_k(x)(a_j\gamma_j(h) - a_k\gamma_k(x))_+$$

$$-\varphi_j(h)\frac{\partial v_j(h)}{\partial y}\sum_k \sum_{x\in D\backslash\{0\}} \varphi_k(x)(a_k\gamma_k(x) - a_j\gamma_j(h))_+ + \varphi_j(h)\varphi(0)a_j\gamma_j(h)\frac{\partial v(0)}{\partial y}.$$

The proof of Corollary A3 is similar to the proof of Theorem 3.

Corollary A4

Under the assumptions of Corollary A3 it follows that

$$\frac{\partial^+ \varphi_j(h)}{\partial w} - \frac{\partial \varphi_j(h)}{\partial w} = \varphi_j(h)\sum_k \sum_{x\in D\backslash\{0\}} \varphi_k(x)\left(\frac{\partial v_k(x)}{\partial y} - \frac{\partial v_j(h)}{\partial y}\right)\min(a_j\gamma_j(h), a_k\gamma_k(x))$$

and

$$\frac{\partial^-\varphi_j(h)}{\partial w}-\frac{\partial\varphi_j(h)}{\partial w}=\varphi_j(h)\sum_k\sum_{x\in D\setminus\{0\}}\varphi_k(x)\left(\frac{\partial v_k(x)}{\partial y}-\frac{\partial v_j(h)}{\partial y}\right)\max(a_j\gamma_j(h),a_k\gamma_k(x))$$

$$+\varphi_j(h)\varphi(0)a_j\gamma_j(h)\left(\frac{\partial v(0)}{\partial y}-\frac{\partial v_j(h)}{\partial y}\right)-\varphi_j(h)a_j\gamma_j(h)\frac{\partial v_j(h)}{\partial y}.$$

The proof of Corollary A4 is similar to the proof of Corollary 1. The next corollary gives the compensated marginal effects of the sectorial probabilities and it follows readily from Corollary A4. By aggregating over all sectors we obtain immediately the next result from Corollary A4.

Corollary A5

Under the assumptions of Corollary A3 we have that

$$\frac{\partial^+\varphi_j^H}{\partial w}=\sum_k\sum_{h\in D\setminus\{0\}}\sum_{x\in D\setminus\{0\}}\varphi_j(h)\varphi_k(x)\frac{\partial v_j(h)}{\partial y}(a_j\gamma_j(h)-a_k\gamma_k(x))_+ +\varphi(0)\sum_{h\in D\setminus\{0\}}\varphi_j(h)\frac{\partial v_j(h)}{\partial y}a_j\gamma_j(h)$$

$$-\sum_k\sum_{h\in D\setminus\{0\}}\sum_{x\in D\setminus\{0\}}\varphi_j(h)\varphi_k(x)\frac{\partial v_k(x)}{\partial y}(a_k\gamma_k(x)-a_j\gamma_j(h))_+$$

and

$$\frac{\partial^-\varphi_j^H}{\partial w}=\sum_k\sum_{h\in D}\sum_{x\in D}\varphi_j(h)\varphi_k(x)\frac{\partial v_k(x)}{\partial y}(a_j\gamma_j(h)-a_k\gamma_k(x))_+ +\varphi(0)\frac{\partial v(0)}{\partial y}\sum_{h\in D}\varphi_j(h)a_j\gamma_j(h)$$

$$-\sum_k\sum_{h\in D}\sum_{x\in D}\varphi_j(h)\varphi_k(x)\frac{\partial v_j(h)}{\partial y}(a_k\gamma_k(x)-a_j\gamma_j(h))_+.$$

From Corollary A5 the next result is immediate.

Corollary A6

Under the assumptions of Corollary A3 we have that

$$\frac{\partial^+\varphi_j^H}{\partial w}-\frac{\partial^-\varphi_j^H}{\partial w}=\varphi(0)\sum_{h\in D}\varphi_j(h)\left(\frac{\partial v_j(h)}{\partial y}-\frac{\partial v(0)}{\partial y}\right)a_j\gamma_j(h).$$

Appendix B.

Proofs

Consider a setting (labor market) with two levels of alternatives, sector and hours of work. Let w_j be the wage rate of sector *j* and y the income and let $\tilde{w}_j$ be the corresponding ex post attribute upper level *r*, *r* = 1, 2, …. Let

$$U_j(h,y) = v_j(h; w_j, y) + \varepsilon_j(h), \quad U(0,y) = v(0,y) + \varepsilon(0)$$

be the ex ante utility of alternative (j,h) and

$$\tilde{U}_j(h,y) = \tilde{v}_j(h; \tilde{w}_j, y) + \varepsilon_j(h) \text{ and } \tilde{U}(0,y) = \tilde{v}(0,y) + \varepsilon(0)$$

the corresponding ex post utilities where the error terms $\{\varepsilon_j(h)\}$ are iid with c.d.f. $\exp(-\exp(-x))$.

Let $Q^H(k,x;j,h)$ be the compensated probability of switching from hours of work *x* in sector *k* to hours of work *h* in sector *j* after a policy intervention given that the ex ante and ex post utility levels are equal. The following Lemma follows from a straight forward extension of the result in Theorem 1.

Lemma B1

Under the assumptions of the multisectoral discrete labor supply model we have that

(B.1) $$Q^H(k,x;j,h) = \int_{y_{jh}}^{y_{kx}} \frac{\exp(v_k(x;w_k,y) + \tilde{v}_j(h;\tilde{w}_j,z))\partial\tilde{v}_j(h;\tilde{w}_j,z)/\partial z}{M(z)^2} \cdot dz$$

and

(B.2) $$Q^H(j,h;j,h) = \frac{\exp(v_j(h;w_j,y))}{M(y_{jh})}$$

for j, k >0, x, h>0, where y_{jh} *and* y_0 *is determined by*

$$v_j(h;w_j,y) = \tilde{v}_j(h;\tilde{w}_j,y_{jh}), \ v(0;,y) = \tilde{v}_0(0;y_0)$$

and

$$M(z) = \sum_{r>0} \sum_{x \in D \setminus \{0\}} \exp(\max(v_r(x;w_r,y), \tilde{v}_r(x;w_r,z))) + \exp(\max(v(0;,y), \tilde{v}(0;,z))).$$

The formulas for $Q^H(0;j,h)$, $Q^H(k,x;0)$ and $Q^H(0;0)$ are similar to the relation in (B.1) and (B.2).

Proof of Theorem A1:

Let $\tilde{w}_j = w_j + \Delta w_j$ where Δw_j is small and positive. Let

$$v_j(h;w_j,y) = u(f(hw_j,y),h) + \log(\theta_j g_j(h)) \quad \text{and } v(0;y) = u(f(0,y),0).$$

Define $y_{jh} = y_{jh}(\tilde{w}_j)$ by

(B.3) $$u(f(w_j h, y), h) = u(f(\tilde{w}_j h, y_{jh}(\tilde{w}_j)), h).$$

Let $Q(k,x;j,h)$ be the probability of switching from hours of work *x* in sector *k* to hours of work *h* in sector *j* after a wage change from w_j to $\tilde{w}_j$ in sector *j*, ceteris paribus. Note that in this case $y_{kx}(w_k) = y$ for $k \neq j$ and $y_{jx}(\tilde{w}_j) < y$. It follows from (B.1) and the mean value theorem for integrals that when $k \neq j$, (B.4) $$Q^H(k,x;j,h) = \frac{(\exp(v_j(x)+v_j(h))\partial v_j(h)/\partial y + O(\Delta w_j))(y - y_{jh}(\tilde{w}_j))}{M(y)^2}$$

$$= -\varphi_j(h)\varphi_k(x)\cdot\frac{\partial v_j(h) y'_{jh}}{\partial y}\cdot\Delta w_j + o(\Delta w_j)$$

where it is understood that the ex post function $\tilde{v}_j$ is equal to the corresponding ex ante function since here there are no changes in taxes and other variables other than wage rates.

$$y'_{jh} = \lim_{\Delta w_j \downarrow 0}\frac{y_{jh}(\tilde{w}_j) - y}{\Delta w_j}.$$

Moreover, from (B.3) it follows by implicit differentiation that

$$0 = \frac{\partial v_j(h;w_j,y_{jh})}{\partial w_j} + \frac{\partial v_j(h;w_j,y_{jh})}{\partial y}\cdot y'_{jh}$$

which gives

(B.5)) $$y'_{jh} = -\frac{\partial v_j(h)/\partial w_j}{\partial v_j(h)/\partial y} = -\frac{f_1'(w_j h, y)h}{f_2'(w_j h, y)} := -\gamma_j(h).$$

When (B.5) is inserted in (B.4) and we obtain that

(B.6) $$Q^H(k,x;j,h) = \varphi_j(h)\varphi_k(x)\gamma_j(h)\frac{\partial v_j(h)}{\partial y}\Delta w_j + o(\Delta w_j)$$

and, similarly, we also have that

(B.7) $$Q^H(0;j,h) = \varphi_j(h)\varphi(0)\gamma_j(h)\frac{\partial v_j(h)}{\partial y}\Delta w_j + o(\Delta w_j).$$

Furthermore, by using the mean value theorem we get from (B.1) and (B.5) that

(B.8) $$Q^H(j,x;j,h) = \varphi_j(h)\varphi_j(x)(\partial v_j(h)/\partial y)(y_{jx}(\tilde{w}_j) - y_{jh}(\tilde{w}_j))_+ + o(\Delta w_j)$$

$$= \varphi_j(x)\varphi_j(h)(\partial v_j(h)/\partial y)(y'_{jx} - y'_{jh})_+\Delta w_j + o(\Delta w_j)$$

$$= \varphi_j(x)\varphi_j(h)(\partial v_j(h)/\partial y)(\gamma_j(h) - \gamma_j(x))_+\Delta w_j + o(\Delta w_j).$$

Let $\tilde{M}(z)$ be the modification of $M(z)$ that consists in replacing w_j by $\tilde{w}_j$. Also, the mean value theorem and (B.2) imply that

(B.9) $$Q^H(j,h;j,h) - \varphi_j(h) = \frac{\exp(v_j(h;w_j,y))}{\tilde{M}(y_{jh})} - \frac{\exp(v_j(h;w_j,y))}{M(y)} = \frac{\varphi_j(h)(M(y) - \tilde{M}(y_{jh}))}{\tilde{M}(y_{jh})}$$

$$= -\varphi_j(h)\cdot\frac{\sum_{x\in D}(\exp(v_j(x;\tilde{w}_j,y_{jh})) - \exp(v_j(x;w_j,y)))_+}{\tilde{M}(y_{jh})}$$

$$= -\varphi_j(h)\sum_{x\in D}\varphi_j(x)\left(\frac{\partial v_j(x;w_j,y)}{\partial w_j} + \frac{\partial v_j(x;w_j,y)}{\partial y}\cdot\frac{\partial y_{jh}}{\partial w_j}\right)_+\Delta w_j + o(\Delta w_j)$$

$$= -\varphi_j(h) \sum_{x \in D} \varphi_j(x) \frac{\partial v_j(x; w_{j,} y)}{\partial y} (\gamma_j(x) - \gamma_j(h))_+ \Delta w_j + o(\Delta w_j)$$

Hence, it follows that

(B.10)

$$\frac{\partial^+ \varphi_j^H(h)}{\partial w_j} = \lim_{\Delta w_j \downarrow 0} \frac{Q^H(0; j,h) + \sum_{k \neq j,0} \sum_{x \in D} Q^H(k,x; j,h) + \sum_{x \in D \setminus \{h\}} Q^H(j,x; j,h) + Q^H(j,h; j,h) - \varphi_j(h)}{\Delta w_j}$$

Hence, (B.6) to (B.10) yield

$$\frac{\partial^+ \varphi_j^H(h)}{\partial w_j} = \sum_{x \in D} \varphi_j(x) \varphi_j(h) \cdot \frac{\partial v_j(h)(\gamma_j(h) - \gamma_j(x))_+}{\partial y} - \varphi_j(h) \sum_{x \in D} \varphi_j(x) \cdot \frac{\partial v_j(x)(\gamma_j(x) - \gamma_j(h))_+}{\partial y}$$

$$+ \varphi_j(h) \sum_{k \neq j} \sum_{x \in D} \varphi_k(x) \cdot \frac{\partial v_j(h) \gamma_j(h)}{\partial y} + \varphi(0) \varphi_j(h) \frac{\partial v_j(h) \gamma_j(h)}{\partial y}$$

which equals the stated result.

Consider next the cross marginal compensated effect implied by a wage rate increase in sector *k*, ceteris paribus.. As above let $\tilde{w}_k = w_k + \Delta w_k$, where Δw_k is small and positive and $k \neq j$. Then it follows that $y_{rx} = y$ when $r \neq k$ and $y_{kx} < y$ which implies that $Q^H(r,x; j,h) = Q^H(0; j,h) = 0$ when $(r,x) \neq (j,h)$. Furthermore, it follows from (B.2) that

$$Q^H(j,h,j,h) = \frac{\exp(v_j(h; w_j, y))}{\sum_{x \in D} \exp(v(x; \tilde{w}_k, y)) + \sum_{r \neq k} \sum_{x \in D} \exp(v_r(x; w_r, y)) + \exp(v(0; y))}.$$

Hence,

$$Q^H(j,h; j,h) - \varphi_j(h) = \frac{\exp(v_j(h; w_j, y)) \sum_{x \in D} [\exp(v_k(x; w_k, y)) - \exp(v_k(x; \tilde{w}_k, y))]}{M(y)^2} + o(\Delta w_k).$$

By the mean value theorem the latter expression implies that

(B.11) $$Q^H(j,h; j,h) - \varphi_j(h) = \frac{\partial \varphi_j(h)}{\partial w_k} \Delta w_k + o(\Delta w_k).$$

Hence, it follows from (B.10) and (B.11) that

$$\frac{\sum_r \sum_{x \in D} Q^H(r,x; j,h) - \varphi_j(h)}{\Delta w_k} = \frac{\sum_{x \in D} Q^H(j,x; j,h) - \varphi_j(h)}{\Delta w_k} = \frac{Q^H(j,h; j,h) - \varphi_j(h)}{\Delta w_k} \underset{\Delta w_k \to 0}{\to} \frac{\partial \varphi_j(h)}{\partial w_k}.$$

Similarly, we have that $Q^H(r,x;0) = 0$ and

$$Q^H(0;0) - \varphi(0) = \frac{\partial \varphi(0)}{\partial w_k} \Delta w_k + o(\Delta w_k).$$

Thus, in this case we conclude that

$$\frac{\partial^+ \varphi_j(h)}{\partial w_k} = \frac{\partial \varphi_j(h)}{\partial w_k} \quad \text{and} \quad \frac{\partial^+ \varphi(0)}{\partial w_k} = \frac{\partial \varphi(0)}{\partial w_k}.$$

Consider next the direct marginal effect when the wage rate in sector *j* decreases, that is, Δw_j is negative. In this case $y_{rx}(w_r) = y$ when $r \neq j$ and $y_{jx}(\tilde{w}_j) > y$ for all *x*. As a result we obtain from (B.1) that $Q^H(r,x,j,h) = 0$ for $r \neq j$, and $Q^H(0;j,h) = 0$. Similarly to (B.8) it follows from (B.1) that

(B.12) $$Q^H(j,x,j,h) = \varphi_j(h)\varphi_j(x)\frac{\partial v_j(h)}{\partial y}(y_{jx}(w_j) - y_{jh}(w_j))_+ + o(\Delta w_j)$$

$$= \varphi_j(h)\varphi_j(x)\frac{\partial v_j(h)}{\partial y}((y'_{jx} - y'_{jh})\Delta w_j)_+ + o(\Delta w_j)$$

$$= \varphi_j(h)\varphi_j(x)\frac{\partial v_j(h)}{\partial y}((\gamma_j(h) - \gamma_j(x))\Delta w_j)_+ + o(\Delta w_j)$$

$$= -\varphi_j(h)\varphi_j(x)\frac{\partial v_j(h)}{\partial y}((\gamma_j(x) - \gamma_j(h))_+\Delta w_j + o(\Delta w_j)$$

and

(B.13) $$Q^H(j,x;0) = -\varphi_j(x)\varphi(0)\frac{\partial v(0)}{\partial y}\gamma_j(x)\Delta w_j + o(\Delta w_j).$$

Furthermore, we have

(B.14) $$Q^H(j,h;j,h) - \varphi_j(h) = \frac{\exp(v_j(h;w_j,y))}{\tilde{M}(y_{jh}(\tilde{w}_j))} - \varphi_j(h)$$

$$= -\varphi_j(h)\frac{\sum_{x\in D\setminus\{0\}}(\exp(v_j(x;\tilde{w}_j,y_{jh}(w_j)) - \exp(v_j(x;w_j,y)))_+ + \sum_{r\neq j}\sum_{x\in D\setminus\{0\}}(\exp(v_r(x;w_r,y_{jh}(w_j)) - \exp(v_r(x;w_r,y)))}{M(y)}$$

$$+\frac{\exp(v(0,y_{jh}(w_j)) - \exp(v(0,y))}{M(y)} + o(\Delta w_j)$$

$$= -\varphi_j(h)\frac{\sum_{x\in D}\exp(v_j(x))\left(\frac{\partial v_j(x)}{\partial w_j}\Delta w_j + \frac{\partial v_j(x)}{\partial y}\cdot\frac{\partial y_{jh}}{\partial w_j}\Delta w_j\right)_+}{M(y)}$$

$$-\frac{\sum_{r\neq j}\sum_{x\in D\setminus\{0\}}\exp(v_r(x))\frac{\Delta w_j\partial v_r(x)}{\partial y}\cdot\frac{\partial y_{jh}}{\partial w_j}}{M(y)} - \frac{\exp(v(0))\frac{\Delta w_j\partial v(0)}{\partial y}\cdot\frac{\partial y_{jh}}{\partial w_j}}{M(y)} + o(\Delta w_j)$$

$$= -\varphi_j(h)\sum_{x\in D}\varphi_j(x)\frac{\partial v_j(x)}{\partial y}(\gamma_j(x)\Delta w_j - \gamma_j(h)\Delta w_j)_+ - \varphi_j(h)\sum_{r\neq j}\sum_{x\in D\setminus\{0\}}\varphi_r(x)\frac{\Delta w_j\partial v_r(x)}{\partial y}\cdot\frac{\partial y_{jh}}{\partial w_j}$$

$$-\varphi_j(h)\varphi(0)\frac{\Delta w_j\partial v(0)}{\partial y}\cdot\frac{\partial y_{jh}}{\partial w_j} + o(\Delta w_j)$$

$$= \Delta w_j\varphi_j(h)\sum_{x\in D}\varphi_j(x)\frac{\partial v_j(x)}{\partial y}(\gamma_j(h) - \gamma_j(x))_+ - \Delta w_j\sum_{r\neq j}\sum_{x\in D\setminus\{0\}}\varphi_r(x)\frac{\partial v_r(x)}{\partial y}\cdot\gamma_j(h) + o(\Delta w_j)$$

$$\Delta w_j \varphi_j(h)\varphi(0)\frac{\partial v(0)}{\partial y}$$

and

(B.15) $Q^H(0;0) - \varphi(0) = 0.$

From (B.13), (B.14) and (B.15) we obtain that

$$\frac{\partial^- \varphi_j^H(h)}{\partial w_j} = -\varphi_j(h)\sum_{x\in D}\varphi_j(x)\frac{\partial v_j(h)}{\partial y}(\gamma_j(x)-\gamma_j(h))_+ + \varphi_j(h)\sum_{x\in D}\varphi_j(x)\frac{\partial v_j(x)}{\partial y}(\gamma_j(h)-\gamma_j(x))_+$$

$$+\varphi_j(h)\gamma_j(h)\sum_{r\neq j,0}\sum_{x\in D}\varphi_r(x)\frac{\partial v_r(x)}{\partial y} + \varphi_j(h)\gamma_j(h)\varphi(0)\frac{\partial v(0)}{\partial y}.$$

Consider finally the cross compensated marginal effects when $\Delta w_k < 0.$ In this case it follows that $y_{rx} = y$ for $r \neq k$ and $y_{kx}(\tilde{w}_k) > y.$ Hence, (B.2) implies that $Q^H(j,h;j,h) = \varphi_j(h).$ Furthermore, (B.1) implies that $Q^H(r,x;j,h) = 0$ for $r \neq k$ and $(r,x) \neq (j,h).$ Hence, using (B.1) we obtain that

$$\sum_r\sum_{x\in D} Q^H(r,x;j,h) - \varphi_k(h) = \sum_{x\in D} Q^H(k,x;j,h)$$

$$= \sum_{x\in D}(\varphi_k(x)\varphi_j(h)\partial v_j(h)/\partial y + O(\Delta w_k))(y_{kx}(\tilde{w}_k) - y))$$

$$= \sum_{x\in D}\varphi_k(x)\varphi_j(h)(\partial v_j(h)/\partial y)y'_{kx}(w_k)\Delta w_k + o(\Delta w_k)$$

$$= -\varphi_j(h)(\partial v_j(h)/\partial y)\sum_{x\in D}\varphi_k(x)\gamma_k(x)\Delta w_k + o(\Delta w_k).$$

When dividing the expression above by Δw_k and letting Δw_k tend towards zero yields

$$\frac{\partial^- \varphi_j(h)}{\partial w_k} = -\varphi_j(h)(\partial v_j(h)/\partial y)\sum_{x\in D}\varphi_k(x)\gamma_k(x).$$

Similarly, it follows that

$$\frac{\partial^- \varphi(0)}{\partial w_k} = -\varphi(0)(\partial v(0)/\partial y)\sum_{x\in D}\varphi_k(x)\gamma_k(x).$$

Q.E.D.

Proof of Theorem 3:

It follows from Theorem A1 that

$$\frac{\partial^+ \varphi_j^H(h)}{\partial w_j} - \frac{\partial \varphi_j(h)}{\partial w_j} = \frac{\partial v_j(h)}{\partial y}\varphi_j(h)\sum_{x\in D}\varphi_j(x)(\gamma_j(h)-\gamma_j(x))_+ - \varphi_j(h)\sum_{x\in D}\varphi_j(x)\frac{\partial v_j(x)}{\partial y}(\gamma_j(x)-\gamma_j(h))_+$$

$$-\frac{\partial v_j(h)}{\partial y}\gamma_j(h)\varphi_j(h) + \varphi_j(h)\sum_{x\in D}\varphi_j(x)\gamma_j(x)\frac{\partial v_j(x)}{\partial y} + \varphi_j(h)(1-\varphi_j)\frac{\partial v_j(h)}{\partial y}\gamma_j(h)$$

$$= -\frac{\partial v_j(h)}{\partial y}\varphi_j(h)\sum_{x\in D}\varphi_j(x)(\gamma_j(h) - (\gamma_j(h)-\gamma_j(x))_+) + \varphi_j(h)\sum_{x\in D}\varphi_j(x)\frac{\partial v_j(x)}{\partial y}(\gamma_j(x) - (\gamma_j(x)-\gamma_j(h))_+).$$

Next, note that

$$\gamma_j(h)-(\gamma_j(h)-\gamma_j(x))_+=\gamma_j(x)-(\gamma_j(x)-\gamma_j(h))_+=\min(\gamma_j(x),\gamma_j(h)).$$

Hence, the above expression reduces to

$$\frac{\partial^+\varphi_j^H(h)}{\partial w_j}-\frac{\partial\varphi_j(h)}{\partial w_j}=\varphi_j(h)\sum_{x\in D}\left(\frac{\partial v_j(x)}{\partial y}-\frac{\partial v_j(h)}{\partial y}\right)\varphi_j(x)\min\big(\gamma_j(h),\gamma_j(x)\big).$$

Similarly,

$$\frac{\partial^-\varphi_j^H(h)}{\partial w_j}-\frac{\partial\varphi_j^H(h)}{\partial w_j}=\frac{\partial v_j(h)}{\partial y}\varphi_j(h)\sum_{x\in D}\varphi_j(x)(\gamma_j(h)-\gamma_j(x))_+-\varphi_j(h)\sum_{x\in D}\varphi_j(x)\frac{\partial v_j(x)}{\partial y}(\gamma_j(x)-\gamma_j(h))_+$$

$$-\frac{\partial v_j(h)}{\partial y}\gamma_j(h)\varphi_j(h)+\varphi_j(h)\sum_{x\in D}\varphi_j(x)\gamma_j(x)\frac{\partial v_j(x)}{\partial y}$$

$$+\varphi_j(h)\gamma_j(h)\sum_{r>0}\sum_{x\in D}\frac{\partial v_r(x)}{\partial y}\varphi_r(x)-\varphi_j(h)\gamma_j(h)\sum_{x\in D}\frac{\partial v_j(x)}{\partial y}\varphi_j(x)+\varphi_j(h)\gamma_j(h)\varphi(0)\frac{\partial v(0)}{\partial y}$$

$$=\varphi_j(h)\sum_{x\in D}\varphi_j(x)\frac{\partial v_j(x)}{\partial y}(\gamma_j(x)-(\gamma_j(x)-\gamma_j(h))_+)-\varphi_j(h)\sum_{x\in D}\varphi_j(x)\frac{\partial v_j(h)}{\partial y}(\gamma_j(h)-(\gamma_j(h)-\gamma_j(x))_+)$$

$$+\frac{\partial v_j(h)}{\partial y}\gamma_j(h)\varphi_j(h)(\varphi_j-\varphi_j(h))+\varphi_j(h)\gamma_j(h)\sum_{r\neq j,0}\sum_{x\in D}\frac{\partial v_r(x)}{\partial y}\varphi_r(x)+\varphi_j(h)\gamma_j(h)\varphi(0)\frac{\partial v(0)}{\partial y}$$

$$=\varphi_j(h)\sum_{x\in D}\varphi_j(x)\left(\frac{\partial v_j(x)}{\partial y}-\frac{\partial v_j(h)}{\partial y}\right)\max(\gamma_j(x),\gamma_j(h))\ +\gamma_j(h)\varphi_j(h)\left(\sum_r\sum_{x\in D}\frac{\partial v_r(x)}{\partial y}\varphi_r(x)\right).$$

Finally, the formula for

$$\frac{\partial^-\varphi_j^H(h)}{\partial w_k}-\frac{\partial\varphi_j(h)}{\partial w_k}$$

is straight forward to verify.

Q.E.D.

Proof of Corollary A1:

From Theorem 3 it follows that

$$\sum_h\left(\frac{\partial^+h\varphi_j^H(h)}{\partial w_j}-\frac{\partial h\varphi_j(h)}{\partial w_j}\right)=\sum_h\sum_x h\varphi_j(h)\frac{\partial v_j(x)}{\partial y}\min\big(\gamma_j(h),\gamma_j(x)\big)\varphi_j(x)$$

$$-\sum_h\sum_x h\varphi_j(h)\frac{\partial v_j(h)}{\partial y}\min\big(\gamma_j(h),\gamma_j(x)\big)\varphi_j(x)$$

$$=-E\left(-\frac{\partial v_j(\tilde{h}_j')}{\partial y}\tilde{h}_j\min\big(\gamma_j(\tilde{h}_j),\gamma_j(h_j')\big)\right)+E\left(-\frac{\partial v_j(\tilde{h}_j)}{\partial y}\tilde{h}_j\min\big(\gamma_j(\tilde{h}_j),\gamma_j(h_j')\big)\right)$$

$$= Cov\left(\tilde{h}_j, -\frac{\partial v_j(\tilde{h}'_j)}{\partial y}\min\left(\gamma_j(\tilde{h}_j), \gamma_j(h'_j)\right)\right) - Cov\left(\tilde{h}_j, -\frac{\partial v_j(\tilde{h}_j)}{\partial y}\min\left(\gamma_j(\tilde{h}_j), \gamma_j(h'_j)\right)\right).$$

The remaining marginal effects given in Corollary A1 are proved in a similar way.

Q.E.D.

Proof of Corollary A2:

Recall that $-\partial v_j(h)/\partial y$ is increasing in *h*. The covariance

$$Cov\left(\tilde{h}_j, -\frac{\partial v_j(\tilde{h}_j)}{\partial y}\min\left(\gamma_j(\tilde{h}_j), \gamma_j(h'_j)\right)\right)$$

is therefore positive and greater than the covariance

$$Cov\left(\tilde{h}_j, -\frac{\partial v_j(\tilde{h}'_j)}{\partial y}\min\left(\gamma_j(\tilde{h}_j), \gamma_j(\tilde{h}'_j)\right)\right)$$

because in the latter expression $\partial v_j(\tilde{h}'_j)/\partial y$ is independent of $\tilde{h}_j$. Hence, the result of Corollary A2 follows from Corollary A1.

Q.E.D.

Proof of Theorem 4:

For simplicity and with no essential lack of generality we only consider the case with only one observable type of agents and one sector. Consider a reform where the tax parameter changes from t to $\tilde{t}$. Let $\varphi(h)$ and $\tilde{\varphi}(h,z)$ denote the ex ante and ex post probability of working *h* hours with ex post non-labor income equal to z. Furthermore, y_h is determined by $u(hw - \tilde{T}(h) + y_h, h) = u(hw - T(h) + y)$ where T and $\tilde{T}$ represent the ex ante and ex post tax system and y is the ex ante income. Then it follows from Dagsvik and Karlström (2005) that

(B.17) $$E(Y(t,\tilde{t},y) \mid y) = \sum_{h\in D}\int_0^{y_h}\tilde{\varphi}(h,z)dz.$$

Since the choice probabilities add up to one, it follows from (B.17) that

(B.18) $$E(Y(t,\tilde{t},y) \mid y) - y = \sum_{h\in D}\int_0^{y_h}\tilde{\varphi}(h,z)dz - y = \sum_{h\in D}\left(\int_0^{y_h}\tilde{\varphi}(h,z)dz - \int_0^{y}\varphi(h)dz\right)$$

$$= \sum_{h\in D}\left(\int_0^{y_h}\tilde{\varphi}(h,z)dz - \int_0^{y}\tilde{\varphi}(h,z)dz + \int_0^{y}(\tilde{\varphi}(h,z) - \varphi(h))dz\right)$$

$$= \sum_{h\in D}\left(\int_0^{y_h}\tilde{\varphi}(h,z)dz - \int_0^{y}\tilde{\varphi}(h,z)dz\right) + \int_0^{y}\sum_{h\in D}(\tilde{\varphi}(h,z) - \varphi(y))dz = \sum_{h\in D}\left(\int_0^{y_h}\tilde{\varphi}(h,z)dz - \int_0^{y}\tilde{\varphi}(h,z)dz\right).$$

Due to the mean value theorem the last equation implies that

$$E(Y(t,\tilde{t},y)\mid y)-y=\sum_{h\in D}\int_{y}^{y_h}\tilde{\varphi}(h,z)dz=\sum_{h\in D}(y_h-y)\varphi(h)+o(\tilde{t}-t)$$

which implies that

$$\lim_{\tilde{t}\to t}\frac{E(Y(t,\tilde{t},y)\mid y)-y}{\tilde{t}-t}=\sum_{h\in D}\varphi(h)\frac{\partial y_h}{\partial t}.$$

From the definition of y_h given above it follows by implicit differentiation that $\partial y_h/\partial t=\partial T(h)/\partial t$. This completes the proof.