# LABOR SUPPLY AS A CHOICE AMONG LATENT JOBS: UNOBSERVED HETEROGENEITY AND IDENTIFICATION 

JOHN K. DAGSVIK ${ }^{\text {a,b }}$ AND ZHIYANG JIA ${ }^{\text {a* }}$<br>${ }^{a}$ Research Department, Statistics Norway, Oslo, Norway<br>${ }^{\text {b }}$ Frisch Centre for Economic Research, Oslo, Norway


#### Abstract

SUMMARY This paper discusses aspects of a modeling framework in which the notion of job choice is fundamental. In this framework, workers are assumed to have preferences over latent job opportunities belonging to worker-specific choice sets from which they choose their preferred job. The main contribution of this paper is an analysis of the identification problem under various conditions, when conventional cross-section micro-data are applied. The modeling framework is applied to analyze labor supply behavior for married/cohabiting couples using Norwegian micro data. Based on the empirical model, we discuss further qualitative properties of the model and simulation of counterfactual policy reforms. Copyright © 2015 John Wiley \& Sons, Ltd.


Received 1 October 2012; Revised 23 September 2014


Supporting information may be found in the online version of this article

## 1. INTRODUCTION

In the traditional approach to labor supply modelling, individual behavior is viewed as a choice among feasible leisure and disposable income combinations. This approach has been criticized for ignoring important behavioral aspects, namely that individuals in the labor market typically have preferences over job types and may face restrictions on their choices regarding job opportunities and hours of work. Recently, the discrete-choice approach to labor supply modeling has gained widespread popularity, mainly because it is much more practical than the traditional continuous approach based on marginal calculus (See Bloemen and Kapteyn, 2008). However, from a theoretical perspective, the conventional discrete-choice approach represents no essential departure from the traditional approach. This is because the only new assumptions made are that the set of feasible hours of work is finite and the random components of the utility function have particular distributional properties. ${ }^{1}$

The purpose of this paper is to discuss identification and other aspects of an extended version of the conventional discrete-choice model (latent job choice model) that allow for agents' preferences being dependent on non-pecuniary job attributes, as well as allowing for possible restrictions on hours of work opportunities. In the latent job choice model, originally proposed by Dagsvik (1994), the starting point is the assumption that a worker's labor supply follows from his or her job choice. More precisely, labor supply is viewed as resulting from a choice among latent job 'packages', each of which is characterized by an offered wage rate, offered hours of work and

[^0]non-pecuniary (qualitative) attributes describing the nature of the job-specific tasks to be performed. This setup can be viewed as a version of Lancaster's characteristic approach-see Lancaster (1966, 1971)-where agents have preferences regarding not only consumption and leisure but also job attributes. The characteristic approach is intuitively appealing because it shifts the focus to qualitative aspects of the labor market that everyday life experiences tell us are important. Examples of such latent job attributes of major importance are job-specific tasks to be performed, location and quality of the social and physical environment. More recently, Farzin (2009) has discussed the effects of including non-pecuniary variables explicitly in the traditional labor supply model. He argues that ignoring such aspects of the jobs can result in biased estimates and thus lead to misleading policy proposals. Further related approaches are put forward by Sattinger $(1993,1995)$ and van Ophem et al. (1993).

The latent job choice model allows us to address neglected aspects in traditional labor supply analysis: namely that workers face important restrictions on their job choice in the labor market (Dagsvik et al., 2014). Versions of the job choice model have been introduced and applied before: see Aaberge et al. (1995, 1999), Dagsvik and Strøm (2004, 2006), Dagsvik and Jia (2006), Kornstad and Thoresen (2007), Di Tommaso et al. (2009), Dagsvik et al. (2011, 2014).

This paper contains a number of new contributions. First, we analyze the identification problem in depth. The identification problem in this type of model differs from standard identification results of discrete-choice models because the present model contains representations of both preferences and choice constraints. We also extend the identification analysis to the case where the distribution of offered wage rates depends on unobserved individual characteristics. The empirical literature on job choice models cited above differs greatly in this respect. Aaberge et al. (1995, 1999) assume that there is no unobserved variation in wages across workers, so that any unobserved heterogeneity in wages is due to firm characteristics. In contrast, Dagsvik and Strøm (2006), Dagsvik and Jia (2006) and Dagsvik et al. (2011) assume that wage variation is due solely to person-specific characteristics. In this paper, we clarify the differences between these two approaches and compare their empirical performances using a sample of Norwegian married/cohabiting couples. Finally, based on the model estimates, we discuss properties of the respective models and the issue of policy simulations. In particular, we show how one can simulate the effect of changes in restrictions on labor supply.

The paper is organized as follows. In Section 2 we discuss the basic structure of the modeling framework. In Section 3 we consider identification issues. In Section 4 we first report results from an empirical application. Finally, we discuss how to simulate the effect of a particular reform in the restriction on working hours.

## 2. THE MODELING FRAMEWORK

Let $U(C, h, z)$ be the (ordinal) utility function of the household, where $C$ denotes household consumption (disposable income), $h$ is hours of work, $z=1,2, \ldots$, indexes market opportunities (jobs) and $z=-1,-2, \ldots$, indexes non-market opportunities. For a market opportunity $z$, associated hours of work and wage rate are assumed fixed and equal to $(H(z), W(z))$, where $H(z), z=1,2, \ldots$, takes the value on a set $D$ and $W(z)$ is positive. When $z$ is negative $H(z)=0$. In addition to the economic budget constraint, there are restrictions on the set of available market opportunities faced by a specific worker. This is because there are job types for which the worker is not qualified and there may be variations in the set of job opportunities for which he or she is qualified. In addition, due to competition in the labor market, jobs for which a worker is qualified may not necessarily be available to him or her. However, the choice sets of market and non-market opportunities are unobserved by the researcher.

Assumption 1. The utility function has the structure

$$
\begin{equation*}
U(C, h, z)=v(C, h) \varepsilon(z) \tag{1}
\end{equation*}
$$

for $z=\ldots,-2,-1,1,2, \ldots$, where $v(\cdot)$ is a positive deterministic function defined on $\left[C_{0}, \infty\right) \times[0, M]$, $\varepsilon(z)$ is a positive random taste shifter, $C_{0}>=0$ is a known constant that represents subsistence consumption and $M$ is the maximum hours of work.

The random taste shifters are supposed to capture the effect of unobserved heterogeneity in preferences over non-pecuniary attributes that affect preferences across agent and across alternatives. Whereas the functional form of the deterministic part of the utility, $v(C, h)$, can be very general, the separability condition in Assumption 1 is crucial. It may not hold in general because the error terms may depend systematically on disposable income and hours of work.

For given hours and wage rate, $h$ and $w$, the economic budget constraint is represented by $C=f(h w, I)$, where $I$ is non-labor income and $f(\cdot)$ is the net of tax function that transforms gross household income into after-tax household income. The function $f(\cdot)$ can in principle capture all details of the tax and benefit system.

The next assumption concerns the representation of choice sets. Here, choice sets are allowed to be random in order to accommodate that they may vary across observable identical agents due to unobserved heterogeneity in opportunities. For technical convenience, we assume at the outset that the choice sets may contain an infinite number of job and non-market opportunities.

Assumption 2. The taste shifters $\{\varepsilon(z), z=\ldots,-2,-1\}$ associated with the available non-market opportunities and $\{\varepsilon(z), z=1,2, \ldots\}$ associated with the available jobs are realizations of two independent Poisson processes on $(0, \infty)$. The intensity of the non-market process is equal to $\varepsilon^{-2}$ and the intensity of the market process is equal to $\theta \varepsilon^{-2}$, where $\theta$ is a positive constant. The available (offered) combinations of job-specific hours of work and wage rates $\{(H(z), W(z)), z=1,2, \ldots\}$ are independent of the corresponding taste shifters and are distributed on $D \times(0, \infty)$, where $D$ is finite and contains at least three points, according to a joint probability density function (p.d.f.) $g_{1}(h) g_{2}(w \mid h)$.

Assumption 2 asserts that the taste shifters associated with the set of available opportunities are independently scattered on the positive part of the real line but in a non-homogeneous way. The probability that there is a job $z$ in the choice set with taste shifter $\varepsilon(z) \in(\varepsilon, \varepsilon+\Delta \varepsilon)$ for some given positive $\varepsilon$ is (approximately) equal to $\theta \varepsilon^{-2} \Delta \varepsilon$. Since $\theta \varepsilon^{-2}$ is decreasing in $\varepsilon$, it means that the intensity is large for $\varepsilon$ close to zero. In other words, while many available jobs are unattractive, there are relatively few attractive jobs available. For example, jobs and non-market opportunities may be located in another region or country.

The parameter $\theta$ is clearly a measure of job availability since a high value of $\theta$ means that there is a high probability that a job with a given level of $\varepsilon(z)$ is available. Dagsvik (1994) has demonstrated that $\theta$ can be interpreted as the ratio of the number of available market opportunities of interest (to the agent) to the number of non-market opportunities of interest. Note also that whereas in Assumption 1 preferences of hours are defined on the continuous set $[0, M]$, the set of feasible hours of work, $D$, is a discrete subset of $(0, M)$. The interpretation of $g_{1}(h)$ is as the probability that a job $z$ with $H(z)$ $=h$ is available to the agent. The interpretation of $g_{2}(w \mid h) \Delta w$ is as the probability (when $\Delta w$ is small) that a job $z$ with $W(z) \in(w, w+\Delta w)$ is available to the agent, given that $H(z)=h$. In the following we shall sometimes call $\theta g_{1}(h) g_{2}(w \mid h)$ the opportunity measure and $g_{1}(h) g_{2}(w \mid h)$ the opportunity density. A motivation for this particular type of representation of the set of available opportunities is given by Dagsvik (1994). He demonstrates that the intensities of the Poisson processes must have the form given in Assumption 2 in order for the choice of job to satisfy the independence from irrelevant alternatives
(IIA) property. ${ }^{2}$ In general, the offered distribution of wage rates may depend on hours of work, similarly to Moffitt (1984). Aaronson and French (2009) have given a theoretical argument that supports this possibility.
Let $\varphi(h, w \mid I)$ denote the joint density of hours of work and wage rate of the chosen job, given nonlabor income $I$, and similarly let $\varphi(0,0 \mid I)$ be the probability of not working. ${ }^{3}$

Theorem 1. If Assumptions 1 and 2 hold, then the joint p.d.f. of observing hours and wage rate combination $(h, w)$ is given by

$$
\begin{equation*}
\varphi(h, w \mid I)=\frac{v(f(h w, I), h) \theta g_{1}(h) g_{2}(w \mid h)}{v(f(0, I), 0)+\theta \sum_{r \in D} \int_{0}^{\infty} v(f(r y, I), r) g_{1}(r) g_{2}(y \mid r) \mathrm{d} y} \tag{2a}
\end{equation*}
$$

for $h>0$, and the probability of not working is given by

$$
\begin{equation*}
\varphi(0,0 \mid I)=\frac{v(f(0, I), 0)}{v(f(0, I), 0)+\theta \sum_{r \in D} \int_{0}^{\infty} v(f(r y, I), r) g_{1}(r) g_{2}(y \mid r) \mathrm{d} y} \tag{2b}
\end{equation*}
$$

In the case with discrete conditional distribution of offered wage rates (or continuous distribution of offered hours of work) similar expressions as in equation ( $2 \mathrm{a}, \mathrm{b}$ ) follow. The only difference is that the integration is replaced by summation (or summation is replaced by integration). The proof of Theorem 1 follows from Dagsvik (1994) but for the reader's convenience a simplified version is given in the supplementary Appendix (supporting information). This proof also covers the special case with degenerate distribution of offered wage rates.
It may be the case that many people view the number of available market opportunities of interest to be less than the number of nonmarket opportunities of interest. In addition, since $\theta$ enters the model as a factor multiplied by $v(C, h)$ it means that $\theta$ also can capture psychological 'costs' of working. For these reasons one might expect that in empirical applications $\theta$ will be less than one.

We noted above that the model in Theorem 1 is consistent with an interpretation with stochastic sets of available job opportunities. There may, however, be additional unobserved heterogeneity that could enter the opportunity measure in other ways. Note that $g_{2}(w \mid h)$ may be interpreted as the conditional distribution of offered wages given offered hours of work equal to $h$. The model ( $2 \mathrm{a}, \mathrm{b}$ ) above assumes that $g_{2}(w \mid h)$ is the same for observable identical individuals (Assumption 2). This may be rather restrictive.
Studies on wage formations and wage dispersion have stressed the importance of unobserved individual heterogeneity; see, for example, Abowd et al. (1999) and Mortensen (2003). We will next consider a typical approach: namely that the distribution of offered wage rates for an individual depends both on observed covariates and on unobserved variables summarized in a random-effect component $\eta$, which is assumed to be independent of the taste shifters $\{\varepsilon(z)\}$. It is also possible to allow $\theta$ to depend on $\eta$, i.e. $\theta=\theta(\eta)$. We shall only consider the special case where $g_{2}(w \mid h ; \eta)$ depends on $\eta$ whereas $g_{1}(h)$ is independent of $\eta$. The motivation for this is that we believe that hours restrictions are

[^1]determined to a large extent by institutional regulations determined by negotiation between the unions of employers and workers and therefore not dependent on individual characteristics. The extended choice model is given as
\[

$$
\begin{equation*}
\varphi(h, w \mid I)=E_{\eta}\left(\frac{v(f(h w, I), h) \theta(\eta) g_{1}(h) g_{2}(w \mid h ; \eta)}{v(f(0, I), 0)+\sum_{r \in D} \int_{0}^{\infty} v(f(r y, I), r) \theta(\eta) g_{1}(r) g_{2}(y \mid r ; \eta) \mathrm{d} y}\right) \tag{3a}
\end{equation*}
$$

\]

for $h>0$, and

$$
\begin{equation*}
\varphi(0,0 \mid I)=E_{\eta}\left(\frac{v(f(0, I), 0)}{v(f(0, I), 0)+\sum_{r \in D} \int_{0}^{\infty} v(f(r y, I), r) \theta(\eta) g_{1}(r) g_{2}(y \mid r ; \eta) \mathrm{d} y}\right) \tag{3b}
\end{equation*}
$$

for $h=0$. Whereas model ( $2 \mathrm{a}, \mathrm{b}$ ) implies that the choice probabilities satisfy the IIA property, this is not the case with the model in equation (3a, b) owing to the random effect $\eta$. In the specification of the model in equation (3a, b), the conditional distribution of offered wage rates $g_{2}(w \mid h ; \eta)$ can be represented by an offered wage rate equation which depends on $(H(z), \xi(z), \eta)$, where the latent variable $\xi(z)$ represents jobspecific unobserved variation in offered wage rates across jobs given offered hours of work. A special version of the offered wage rate equation above have an additive separable form such as

$$
\begin{equation*}
\log W(z)=\alpha+\psi(H(z))+\eta+\xi(z) \tag{4}
\end{equation*}
$$

where $\alpha$ is a function of individual characteristics, $\eta, \xi(z)$ and $H(z)$ are assumed to be mutually independent and $\psi(H(z))$ represents the influence of the job-specific offered hours.

The case where equation (4) holds and $\xi(z)=0$ needs further discussion. This corresponds to a model with degenerate distribution of offered wage rates, where the offered wage rates for a given individual vary only with respect to offered hours of work. Examples of studies within the conventional framework of this type include Moffitt (1984), Lundberg (1985), Biddle and Zarkin (1989), Wolf (2002) and Aaronson and French (2004). These authors typically find a positive or inverse U shape relationship between the wage rates and hours worked. This feature is consistent with findings in the literature where part-time workers have been found to earn less than full-time workers. However, the gap seems to be very small, once important job characteristics such as occupations are considered (Manning and Petrongolo, 2008).

Within our framework, when $\xi(z)=0$ it follows from equation (4) that for any job $z$ we have $\eta=\log W(z)-\psi(H(z))-\alpha$. Thus, given that the chosen hours of work and wage rate combination equals ( $h, w$ ), it follows that any job $z$ (say) with hours of work $H(z)$ has wage rate given by $W(z)=w \exp (\psi(H(z))-\psi(h))$. Consequently, in the case with $\xi(z)=0$ equation (3a) reduces to

$$
\begin{equation*}
\varphi(h, w \mid I)=\frac{v(f(h w, I), h) \theta(\log w-\psi(h)-\alpha) g_{1}(h) g_{\eta}(\log w-\psi(h)-\alpha) / w}{v(f(0, I), 0)+\theta(\log w-\psi(h)-\alpha) \sum_{r \in D} v\left(f(r w \exp (\psi(r)-\psi(h)-\alpha, I), r) g_{1}(r)\right.} \tag{5}
\end{equation*}
$$

where $g_{\eta}(\cdot)$ is the p.d.f. of $\eta$. If in addition, we assume that $\psi(\cdot)=0$, i.e. the offered wage rate does not depend on offered hours of work, and that $\theta(\eta)=\theta$, equation (5) reduces to

$$
\begin{equation*}
\varphi(h, w \mid I)=\frac{v(f(h w, I), h) \theta g_{1}(h)}{v(f(0, I), 0)+\theta \sum_{r \in D} v(f(r w, I), r) g_{1}(r)} \cdot \frac{g_{\eta}(\log w-\alpha)}{w} \tag{6}
\end{equation*}
$$

whereas in the analogous expression for the probability of not working the random effect must be integrated out. A two-sector model similar to the model given by equation (6) has been applied by Dagsvik and Strøm (2006) to analyze labor supply and sectoral choice.

When $\eta=\psi(h)=0$ we get a model where offered wage rates vary across different job offers (to a given agent). This model version was applied by Aaberge et al. (1995, 1999).

## 3. IDENTIFICATION

We now turn to a discussion on identification of the model. We start with the simpler case where the random effect in the wage rate equation is ruled out, i.e. the model is as presented in Theorem 1. From Theorem 1 it follows that for positive $h$

$$
\begin{equation*}
\varphi(h, w \mid I) / \varphi(0,0 \mid I)=v(f(h w, I), h) \theta g_{1}(h) g_{2}(w \mid h) / v(f(0, I), 0) \tag{7}
\end{equation*}
$$

Since both $\varphi(h, w \mid I)$ and $\varphi(0,0 \mid I)$ are observable, the right-hand side of equation (7) can be identified non-parametrically. However, it remains unclear if one can separately identify $v(C, h)$ and $\theta g_{1}(h) g_{2}(w / h)$. Essentially, the identification problem arises from the fact that observed labor supply behavior is a result of both preferences (utility function) and latent job choice constraints (opportunity measure) in our model.

If available, information on desired hours of work could be used to identify preferences, since job choice constraints are irrelevant in this case. ${ }^{4}$ Subsequently, based on the estimated preferences, one could, with this information, estimate the opportunity measure using data on actual observed behavior. In this way, both the utility function and the opportunity measure can be identified. This approach is. however. not straightforward because the representation of job types in stated preference surveys may not fully correspond to the variety of job offers agents face in real labor markets.
Note that, even if panel data or independent cross-section data are available, it is hard to see how this would help to solve the identification problem in general. If, for example, preference parameters were assumed to be fixed over time, this would enable us to identify changes in the opportunity measure non-parametrically, but not the level.
In what follows we limit our discussion to the situation where only cross-section data are available. That is, we have observations on hours of work, wage, non-labor income and individual characteristics. Assumption 3 summarizes some useful regularity conditions for the deterministic part of the utility function $v(C, h)$ and the net of tax function.

Assumption 3. The deterministic part of the utility function $v(C, h)$ is continuously differentiable in $C$ and the net of tax function $f(u, I)$ is continuous, piecewise differentiable and strictly increasing in $u$ and $I$.

One important property of our setup is that non-labor income enters the utility specification in a particular manner, namely such that it can generate variation in consumption while keeping hours of work and the wage rate constant, and it enters the model only through consumption, not affecting the

[^2]opportunity measure. The next theorem shows to what extent the model can be non-parametrically identified in this case.

Theorem 2. Assume that Assumptions $1-3$ hold. Then, $v(C, h)$ can be expressed as $v(C, h)=$ $\zeta(C)^{r} \lambda^{*}(C, h) \delta(h)$ for $h>0$, where $\zeta(C)$ and $\lambda^{*}(C, h)$ are identified but $r$ is an unknown constant and $\delta(h)$ an unknown function of $h$.

The proof of Theorem 2 is given in APPENDIX A. Theorem 2 shows that, even under the exclusion restriction that non-labor income does not affect the opportunity measure, the model is nonparametrically unidentified. Dagsvik and Strøm (1997) show identification in an analogous model by assuming that preferences are separable in consumption and hours of work and, more importantly, that fixed cost of working is observed. However, such information is rarely available.

Assumption 4. The offered wage rates and the offered hours of work are independently distributed.
Theorem 3. Assume that Assumptions 1-4 hold. Then the distribution of offered hours is identified and $v(C, h)=\lambda(C, h) \delta(h)$, where $\lambda(C, h)$ is identified but $\delta(h)$ is not identified.

The proof is given in APPENDIX A. When offered wage rates and hours of work are correlated, additional restrictions are needed to achieve identification. Note that, for the purpose of simulating solely the effect of counterfactual changes in taxes and wage rates, it is not necessary to identify $\delta(h)$ and $g_{1}(h)$ separately as long as $g_{1}(h)$ is kept fixed. The reason is that the effect of changes in taxed and wages only enters the model through $C$ and the opportunity density of offered wage rates.

One way to obtain full identification is to make parametric functional form assumptions about both $v(C, h)$ and $\theta g_{1}(h) g_{2}(w \mid h)$. Below, we consider one particular parametric utility specification, namely the generalized Box-Cox function, given by

$$
\begin{equation*}
\log v(C, h)=\gamma_{1}\left(C^{\alpha}-1\right) / \alpha+\gamma_{2}\left((1-h / M)^{\beta}-1\right) / \beta+\gamma_{3}\left(C^{\alpha}-1\right)\left((1-h / M)^{\beta}-1\right) / \alpha \beta \tag{8}
\end{equation*}
$$

See Dagsvik and Strøm (2006) and Dagsvik and Røine Hoff (2011) for a theoretical justification based on invariance principles. ${ }^{5}$

Assumption 5. There exists an interval $\left(u_{-}, u_{+}\right)$such that when $u \in\left(u_{-}, u_{+}\right)$, the marginal net of tax rate w.r.t. non-labor income $\partial f(u, I) / \partial I$ does not vary with $u$.

Assumption 5 is not particularly restrictive. In fact, most tax systems satisfy this condition.
Assumption 6. The function $\log v(C, h)$ is a generalized Box-Cox function as given in equation (8). Furthermore, there exist two points $h_{1} \neq h_{2} \in D$ such that $g_{1}\left(h_{1}\right)=g_{1}\left(h_{2}\right)$.

The following theorem summarizes the key identification results when additional functional form assumptions are made for $v(C, h)$ and $g_{1}(h) g_{2}(w \mid h)$.

Theorem 4. Assume that Assumptions 1-3, 5 and 6 hold. Then the model given by equation ( 2 a , b) is identified.

The proof of Theorem 4 is given in the supplementary Appendix. Theorem 4 gives sufficient conditions for the model in equation ( $2 \mathrm{a}, \mathrm{b}$ ) to be fully identified. Assumption 5 may not be necessary but is made here for convenience.

[^3]We shall next consider the more general setting where we allow for unobserved heterogeneity in the offered wage rate, i.e. the model given by equation ( $3 \mathrm{a}, \mathrm{b}$ ). To achieve identification, it is necessary to introduce an exogenous variable $X$ which influences only the opportunity density.

Assumption 7. The offered wage rates are generated by

$$
\begin{equation*}
\log W(z)=X b+a+\eta+\xi(z) \tag{9}
\end{equation*}
$$

where $X$ is a continuous covariate that does not affect preferences and can take any real value, $\xi(z)$ and $\eta$ are zero mean random variables which are independent of each other and independent of the taste shifters, and $a$ and $b$ are unknown parameters. Furthermore, $\theta$ is either a constant or has the structure

$$
\begin{equation*}
\theta=\theta(a+X b+\eta) \tag{10}
\end{equation*}
$$

Assumption 7 asserts that the offered wage rate equation is additively separable in $a+X b, \eta$ and $\xi(z)$. The formulation in equation (9) is a special case of equation (4) with $\psi(\cdot)=0, \alpha=a+X b$. The parameter $a$ may depend on individual characteristics other than $X$. The motivation for the particular formulation of $\theta(\cdot)$ in Assumption 7 is that $a+X b+\eta$ is supposed to represent the effect of observed and unobserved individual ability which may affect the opportunity measure. Here, $\xi(z)$ and $\eta$ may be discrete or continuous random variables and $\xi(z)$ may also be degenerate. In the following we need to extend our notation of $\varphi(h, w \mid I)$ to $\varphi(h, w I I, X)$ to indicate that the latter p.d.f. is conditional on $X$.

Assumption 8. The function $\varphi(h, w \mid I, X) / v(f(h w, I), h))$ satisfies

$$
\int_{-\infty}^{\infty} \int_{0}^{\infty}\left(|X|^{m}|\log w|^{n} \varphi(h, w \mid I, X) / v(f(h w, I), h)\right) \mathrm{d} w \mathrm{~d} X<\infty
$$

for $m, n=0,1,2, h \in D$ and $I>0 .{ }^{6}$

## Theorem 5.

i. If Assumptions $1-4,7$ and 8 hold, then $v(C, h)$ is identified apart from a multiplicative term that may depend on $\mathrm{h}, \theta(\cdot)$ is identified up to a constant and the distribution of offered wage rates conditional on the random effect is identified.
ii. If Assumptions $1-4$, and 6 to 8 hold then the model in equation (3a, b) is identified.

The proof of Theorem 5 is given in the supplementary Appendix. Theorem 5 extends the identification results of Theorems 3 and 4 to the case with a random effect in the offered wage rate distribution. It is worth noting that the independence condition (Assumption 4) is still maintained in Theorem 5 in contrast to Theorem 4, where offered wage rate and hours of work are allowed to be correlated.

Theorems $2-5$ hold also in the case where the distribution of offered wage rates is discrete or the distribution of offered hours of work is continuous. Furthermore, Theorem 5 holds when $X$ is a discrete variable or when the distribution of offered wage rates across jobs is degenerate, that is, where each agent only faces one individual specific wage rate.

[^4]
## 4. AN EMPIRICAL APPLICATION

### 4.1. Model Specification and Estimation Results

In this section we report results from an empirical application based on micro data from the Norwegian Labor Survey 1997. Table I summarizes the data for our empirical analysis. Details about variable definitions and the data, and the specification of a joint labor supply model for married couples, can be found in the supplementary Appendix.

The systematic part of the utility function is assumed to be a generalized Box-Cox functional form, similarly to equation (8). Alternatively, we could have used a flexible polynomial specification similarly to van Soest et al. (2002). Dagsvik and Strøm (2004, 2006) found that the Box-Cox functional form is more or less as flexible as the polynomial specifications, and in contrast to the latter ones it is globally concave. For each spouse, we specify eight feasible annual hours of work alternatives, namely $0,208,624,1040,1456,1950,2340$ and 2600 . The logarithm of the job availability measures $\theta_{F}$ and $\theta_{M}$ are specified as linear functions of length of schooling. The opportunity probability mass functions of offered hours, $g_{1 k}(h), k=F, M$, are uniform except for peaks at full-time and part-time hours. The full-time peak corresponds to 1950 hours annually ( 37.5 hours a week), while the part-time peak corresponds to 1040 hours annually ( 20 hours a week). The part-time and full-time peaks in the hours distribution are supposed to capture the effect of institutional regulations on hours of work. Recall that this specification of the opportunity distribution is formally equivalent to introducing suitable dummy variables at full-time and part-time hours of work in the utility specification of the conventional discrete-choice specification; see, for example, van Soest (1995). An advantage with our framework is that it provides a theoretical rationale for introducing such dummies, in contrast to the conventional discrete-choice labor supply model.

A central issue in our empirical illustration is how to specify the opportunity density of offered wage rates. A fairly general class of specifications accounting for unobserved heterogeneity across both jobs and agents, as well as allowing for correlation between offered hours of work and offered wage rates, is given in equation (4). It remains, however, to prove that the model is identified in this case. Even in the case where offered hours and wage rates are independent, it is a demanding task to obtain structural

Table I. Summary statistics

|  | Both working |  | Only husband working |  | Only wife working |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Mean | SD | Mean | SD | Mean | SD |
| Men |  |  |  |  |  |  |
| Age | 45.06 | 8.44 | 44.06 | 9.41 | 50.40 | 9.91 |
| Education | 12.58 | 2.89 | 12.26 | 2.71 | 11.80 | 3.03 |
| Experience | 25.48 | 9.14 | 24.80 | 9.99 | 31.60 | 11.08 |
| non-labor income | 6320 | 12032 | 10796 | 15543 | 30604 | 30395 |
| wage rate | 153.82 | 52.98 | 169.11 | 65.47 |  |  |
| weekly hours of work | 38.43 | 5.38 | 39.16 | 5.91 |  |  |
| Women |  |  |  |  |  |  |
| Age | 42.75 | 8.39 | 41.68 | 9.70 | 47.40 | 7.86 |
| Education | 12.10 | 2.72 | 10.87 | 2.26 | 11.80 | 2.86 |
| Experience | 23.65 | 9.33 | 23.81 | 10.71 | 28.60 | 10.01 |
| Non-labor income | 17655 | 16558 | 27689 | 21701 | 13848 | 14219 |
| Wage rate | 120.12 | 37.79 |  |  | 22.82 | 17.98 |
| Weekly hours of work | 30.45 | 8.93 |  |  | 29.78 | 12.06 |
| No. of children 0-6 | 0.37 | 0.68 | 0.70 | 0.89 | 0.20 | 0.45 |
| No. of children 7-18 | 0.85 | 0.97 | 0.94 | 0.97 | 1.00 | 0.71 |
| No. of households | 2254 |  | 256 |  | 5 |  |

estimates due to measurement error in reported hours worked. In our dataset, only weekly hours of work are reported but not the usual number of weeks worked during a year. Furthermore, overtime is not reported. This type of measurement error is common in datasets used for labor supply analysis. See, for example, Borjas (1980) and Blundell et al. (2007) for extensive discussions on how to deal with this problem in conventional labor supply models. This measurement error creates a spurious negative correlation between the observed wage rate and hours in our data ( -0.22 for married women, -0.17 for married men). ${ }^{7}$ This so-called 'division bias' problem poses considerable challenges for empirical analyses based on this type of data.

When offered wage rates and offered hours of work are independent, the division bias problem can be reduced by adopting the three-stage estimation procedure proposed by Dagsvik and Strøm (2004, 2006). This amounts to estimating a reduced-form participation probability in the first stage, and subsequently estimating wage rate equations using the results from the first stage to control for selectivity bias. In the third stage the labor supply model is estimated by the maximum likelihood method after inserting the wage rates predicted by the estimated wage equations into the model and integrating out the respective error terms. Under standard assumptions about the measurement error this leads to unbiased wage rate equations in the second stage apart from the estimate of the variance of the error terms which may be upward biased. This procedure does not, of course, eliminate the measurement error in hours of work but it removes the spurious negative correlation between observed hours and wage rates. Measurement error may still be a problem in the last-stage maximum likelihood estimation, since it may cause misclassification of the dependent variable (individual's working hours). Hausman et al. (1998) suggest a method to control for the misspecification problem in a binary-choice model. However, their method is not readily applicable in our analysis.

Given Assumption 4, we have estimated two special cases of the model in equation (3a, b): Model 1 and Model 2. Using the notation of equation (4), Model 1 is based on the assumption that all observable identical individuals are assumed to face the same offered wage rate distribution across jobs $(\eta=\psi(\cdot)=0)$, whereas in Model 2 each individual faces only one offered wage, though this wage may be different for observationally identical individuals $(\xi(z)=\psi(\cdot)=0)$. Thus the interpretation of the first stage wage rate equations is rather different in the two model versions, although the estimated wage equations have the same structure. In Model 1, the error terms in the wage rate equations measure the variation of wage rate offers across different jobs. In contrast, in Model 2, the error terms represent unobserved individual heterogeneity. ${ }^{8}$

The specification of the wage rate equations is conventional. In Model $1 \xi(z)$ is assumed to be normally distributed with zero mean and in Model $2 \eta$ is assumed to be normally distributed with zero mean. Thus the estimated residual variances in the wage rate equations are interpreted as the variance of $\xi(z)$ and $\eta$ in Models 1 and 2, respectively. The term $\alpha$ is specified as a linear function of length of schooling, potential experience, potential experience squared and a dummy variable for marriage status. Potential experience is defined as age minus years of schooling minus 7. The estimates of the wage rate equations are presented in Table S 1 in the supplementary Appendix. As shown in Table S 1 , the selection bias in the wage rate equations is negligible.

The estimates of the two models are reported in Table S 2 in the supplementary Appendix. In both models the results imply that $\log v(C, h)$ is strictly increasing and concave in consumption and leisure. The marginal utility of leisure of married women is decreasing until about the age of 35 and thereafter

[^5]increasing. The similar pattern holds for the married men, only with the minimum value obtained around the age of 37 . The number of children has a significant effect on the marginal utility of leisure for married women. In contrast, the marginal utility of leisure for married men does not depend significantly on the number of children. This indicates that the female takes more responsibility for children within the family than the male, which is not a surprising result. The measure of the job availability for the married women, $\theta_{F}$, depends positively on the length of schooling $(S)$. Higher educational level increases the job opportunities for married women. The corresponding estimate for married men turns out not to be significant. This may be due to the fact that in our dataset there are only very few married men who are out of the labor force. For both genders, the value of $\theta$ is estimated to be less than 1 . Similar results are reported by Dagsvik and Strøm (2006), Di Tommaso et al. (2009) and Dagsvik et al. (2011). As discussed earlier, this can be interpreted as indicating that the number of interesting and available jobs is smaller than the number of interesting non-market opportunities. Note also that the full-time peak in the opportunity probability mass function of hours for married men is substantially higher than the corresponding peak for married women whereas the opposite is true for the part-time peaks. The reason for this is that women seem to have stronger preferences than men for working in particular labor market sectors (such as health care and teaching) and therefore may face different choice restrictions. This could partly be due to differences in gender-specific human capital investments, which are important in shaping the job choice constraints. ${ }^{9}$

Both Model 1 and Model 2 fit the data quite well. Figures S1 and S2 in the supplementary Appendix display the observed and (aggregate) predicted values of participation and hours of work for each spouse based on our model. Model 2 seems to perform slightly better than Model 1. Since our model represents the behavior of couples, it seems more interesting to consider the joint distribution of hours of work rather than just the marginals. Owing to the problem of thin cells, we are forced to consider joint distributions with highly aggregate hours of work intervals and combine selected hours of work intervals taking into account the fact that some hours of work alternatives are chosen by only a few households in our sample. For men, we distinguish only between the cases where they work less than full time ( $0-1456$ hours a year) or work full time and more ( $<=1950$ hours a year). For women, we distinguish between three options (not working, working less than full time and working full time and more). Altogether, our selected aggregation procedure implies just six cells. Figure 1 shows the predicted and observed distributions of hours of work on these six cells. The fit of both models is quite good, with Model 2 performing slightly better than Model 1.

These visual inspections of the marginal and joint distributions of hours of work can be seen as an informal and crude way to check goodness of fit of the models. Several alternative model evaluation methods are used for discrete-choice models in the literature; see, for example, Train (2009). One popular summary measure analogous to the $R^{2}$ measure applied in regression analysis is McFadden's $\rho^{2}$ measure (pseudo $R^{2}$ ); see McFadden (1973). The respective values of the log-likelihood functions are approximately equal to -5309 and -5243 . The corresponding $\rho^{2}$ values for the two models are 0.49 and 0.50 respectively. These results seem to be consistent with the informal check based on Figure 1, and Figures S1 and S2 in the supplementary Appendix.

Another way of measuring goodness of fit uses tests based on comparing predicted probabilities with corresponding observed frequencies on some partitions of the outcome and exogenous variables in the tradition of the Pearson chi-square statistics. We have applied the chi-square test statistics of Andrews (1988a,b). To this end we used the same partition of cells as above (six cells). The test statistic is given as $F=N v_{N}(\widehat{\beta})^{\prime} \hat{V}^{-1} v_{N}(\widehat{\beta})$, where $v_{N}(\widehat{\beta})$ is the vector of differences between the predicted and observed frequencies on the chosen cells, and $\hat{V}$ is the estimated covariance matrix of $v_{N}(\widehat{\beta}) \sqrt{N}$. In our
${ }^{9}$ In a sector-specific model, as in Dagsvik and Strøm (2006), one could obtain explicit sector-specific opportunity measures.


Figure 1. Predicted and observed joint hours of work distributions for couples, 1997
case, the test statistic $F$ is asymptotically chi-square distributed with 5 degrees of freedom. The test statistic for Model 1 is equal to 57.6 and for Model 2 it is equal to 10.4 . The $5 \%$ significance level for the chi-square distribution with 5 degrees of freedom is 11.07 . Thus Model 2 passes Andrew's test at $5 \%$ significance level, while Model 1 is far from passing the test. Thus Model 2 fits the data better than Model 1 and we therefore select Model 2 as our maintained model.

### 4.2. Aggregate Wage Elasticities and 'Labor Supply Curves'

In this section, we present selected gross wage elasticities and figures for what we call labor supply curves. We have calculated elasticities that take into account both the systematic terms and the unobservables in the model. This means that we account for how the mean of the distribution of labor supply is affected by changes in (say) gross wage levels.

There are two types of elasticities reported in the literature; see, for example, Dagsvik and Strøm (2006) and van Soest and Das (2001). One type is the so-called average elasticity, which is simply the average of the individual elasticities calculated from the formulas for expected hours of work or probability of working given the individual observed characteristics. The second is called aggregated elasticity, which is the elasticity of the aggregate (or average) response (e.g. the elasticity of the population mean hours of work).
Table II displays aggregated gross wage elasticities for both models. The standard errors are generated using bootstrap methods with 100 replications. We have also computed the average elasticities for both models: the estimates are quite close to the aggregated elasticities reported in Table II. The unconditional wage elasticity of hours of work is defined as the elasticity of expected number of hours of work with respect to wage. The corresponding elasticity of hours of work conditional on working can be found by subtracting the wage elasticity of the probability of working from the unconditional wage elasticity of hours of work.
In general, both models show that the gross wage elasticities are moderate for married females and very small for married males. Model 2 seems to predict a stronger labor supply response than Model 1,

Table II. Aggregated gross wage elasticities

|  | Model 1 |  |  | Model 2 |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Women's wage | Men's wage | Both wages | Women's wage | Men's wage | Both wages |
| Probability of working, men | $\begin{gathered} -0.006 \\ (0.003) \end{gathered}$ | $\begin{gathered} 0.007 \\ (0.003) \end{gathered}$ | $\begin{gathered} 0.003 \\ (0.001) \end{gathered}$ | $\begin{gathered} -0.007 \\ (0.003) \end{gathered}$ | $\begin{gathered} 0.010 \\ (0.004) \end{gathered}$ | $\begin{gathered} 0.006 \\ (0.003) \end{gathered}$ |
| Probability of working, women | $\begin{gathered} 0.221 \\ (0.015) \end{gathered}$ | $\begin{gathered} -0.145 \\ (0.020) \end{gathered}$ | $\begin{gathered} 0.091 \\ (0.023) \end{gathered}$ | $\begin{gathered} 0.333 \\ (0.022) \end{gathered}$ | $\begin{gathered} -0.165 \\ (0.022) \end{gathered}$ | $\begin{gathered} 0.205 \\ (0.024) \end{gathered}$ |
| Hours of work, men (unconditional) | $\begin{gathered} -0.022 \\ (0.004) \end{gathered}$ | $\begin{gathered} 0.047 \\ (0.007) \end{gathered}$ | $\begin{gathered} 0.028 \\ (0.008) \end{gathered}$ | $\begin{gathered} -0.022 \\ (0.005) \end{gathered}$ | $\begin{gathered} 0.080 \\ (0.009) \end{gathered}$ | $\begin{gathered} 0.062 \\ (0.009) \end{gathered}$ |
| Hours of work, women (unconditional) | $\begin{gathered} 0.405 \\ (0.025) \end{gathered}$ | $\begin{array}{r} -0.268 \\ (0.034) \end{array}$ | $\begin{gathered} 0.151 \\ (0.050) \end{gathered}$ | $\begin{gathered} 0.618 \\ (0.039) \end{gathered}$ | $\begin{gathered} -0.272 \\ (0.038) \end{gathered}$ | $\begin{gathered} 0.381 \\ (0.054) \end{gathered}$ |

Bootstrapped standard errors in parentheses.
though the differences are not big. Using estimates from Model 2, for married females the own-wage elasticity of the probability of working is equal to 0.33 , which means that if the gross wages of married females were to increase by $5 \%$ (say), then the aggregate proportion of married females working would increase by around $1.5 \%$. The cross-wage elasticity for married women is negative and smaller in size than own wage elasticity, as found in many studies. The elasticity of the probability of working for married women with respect to both spouses' gross wage rate is equal to 0.2 . This means that the proportion of married women working would increase by $1 \%$, if both spouses' wage rates were increased by $5 \%$.

Since the model is highly nonlinear, the aggregate elasticities may cover substantial variation in elasticities across different wage levels and individual characteristics. Remember that when we have estimated the model we can compute wage elasticities conditional on given hypothetical wage levels, without using wage equations. To illustrate the nonlinearity in the labor supply response resulting from wage changes, we have plotted what we call expected labor supply curves. These curves show expected hours of work as a function of the wage rates, conditional on other characteristics (non-labor income, age, family size and opportunity measure). We have generated labor supply curves for married women for a typical household, which is constructed using the sample average value of household characteristics given that the household consists of a husband aged 45 , a wife aged 42 and one school-age child. Both parents have 12 years of education. We look at two different scenarios: the husband's wage is high (200 NOK (Norwegian krone) per hour) ${ }^{10}$ or low ( 100 NOK per hour). Figure 2 shows how the expected annual hours of work of married women in both models vary with her expected gross wage level. Both models predict that labor supply elasticities decrease as the women's wage increases. However, at low wage rate levels the labor supply curves are steeper for Model 2 than for Model 1, whereas the difference is small between the two models at high wage rate levels.

### 4.3. Simulation of Changes in the Opportunity Distribution of Offered Hours of Work

Using our framework, we can also simulate the effect on labor supply from changing the opportunity distribution. Since our model is not an equilibrium one, we can only simulate pure supply effects conditional on given job availability measures, ( $\theta_{F}, \theta_{M}$ ), and the distributions of offered hours and wage rates.

[^6]

Figure 2. Married women's expected hours of work in a typical household, by husband's wage. The shaded area shows the $95 \%$ confidence band

In Norway, there is a high proportion of married women in part-time positions. In our sample, more than $35 \%$ of the married women are working between 20 and 30 hours a week, while only around $40 \%$ are working full time ( 37.5 hours a week). Many of those in part-time jobs are employed in the public sector, especially in health care. Whereas Norwegian working environment legislation opens for the possibility of voluntary part-time jobs, an important reason for the high concentration of part-time workers relates to particular institutional regulations in the public health sector. Part-time positions are sometimes the only positions offered by public health care organizations, especially in more rural areas. The proposed reform can be interpreted as a change in the opportunity distribution of hours for women, obtained by removing the part-time peak and increasing the full-time peak (since the part-time jobs are replaced by full-time ones) while keeping the total number of available jobs unchanged (i.e. $\theta_{F}$ is unchanged). See APPENDIX B for details.

Figure 3 displays the results from such a simulation. As we can see, there is a significant decrease in the proportion of married women who choose to work part time, accompanied by an increase in the proportion of full-time hours of work of about a similar magnitude. In addition, we found that the corresponding labor supply of married men changes very little with the introduction of this reform.

However, one needs to be careful when interpreting this result in the context of reforms specific to labor market sectors (such as the health care sector), since sector-specific preferences and restrictions are not explicitly accounted for in the model. Nevertheless, this simulation exercise clearly illustrates the advantage and potential of our modeling framework.


Figure 3. Hours of work for married women, before and after reform. The $95 \%$ confidence intervals are represented by the error bars

## 5. CONCLUSIONS

The traditional models of labor supply, being versions of the theory of consumer demand with two goods-disposable income and leisure-simplify the choice setting in labor markets. In real labor markets, agents have preferences over pecuniary as well as non-pecuniary aspects of jobs and face limited sets of job opportunities in the labor market due to competition between workers and restrictions resulting from institutional regulations on hours of work.

An essential feature of the modeling framework discussed in this paper is that it allows the researcher to accommodate restrictions on hours of work and the set of feasible jobs typically observed in many datasets. A major contribution of this paper is the analysis of identification. The standard identification results for multinomial and mixed logit models do not apply because our modeling framework contains a representation of both preferences and latent choice constraints (opportunity measure). Although the model is, in general, not identified we have demonstrated how it will be identified under specific conditions even in the presence of unobserved heterogeneity in the offered wage rate distribution.

We have carried out an empirical application based on micro data from Norway. In contrast to Dagsvik and Strøm (2006), who estimated a similar model for married women given the husband's labor supply, we study the joint labor supply behavior for married couples. We have, moreover, estimated two versions of the model based on two 'extreme' assumptions about wage heterogeneity. A model with solely inter-individual variation in wage rates seems to fit the data better than a model that allows only for variation in wage rates across jobs. We have computed labor supply curves to illustrate the substantial nonlinearity in the labor supply responses, as a function of wage rates and to illustrate how these curves depend on the two model formulations. Subsequently, we have shown how one can use the model to simulate the effect of hypothetical changes in the opportunity measure. Changes of this sort cannot be studied using conventional discrete labor supply models.

The data we have applied are not ideal owing to measurement error in the hours of work observations. We are currently working on establishing another dataset with much less measurement error in
hours, which hopefully will be helpful for studying the relationship between the offered wage rates and offered hours within a labor supply modeling framework.

## ACKNOWLEDGEMENTS

We thank J. Aasness, T. O. Thoresen and T. Skjerpen for helpful comments. We are particularly grateful for the extensive and constructive criticism, help and suggestions by T. Magnac and two anonymous referees.

## REFERENCES

Aaberge R, Dagsvik JK, Strøm S. 1995. Labor supply responses and welfare effects of tax reforms. Scandinavian Journal of Economics 97: 635-659.
Aaberge R, Columbino U, Strøm S. 1999. Labor supply in Italy: an empirical analysis of joint household decisions with taxes and quantity constraints. Journal of Applied Econometrics 14: 403-422.
Aaronson D, French E. 2004. The effect of part-time work on wages: evidence from the social security rules. Journal of Labor Economics 22: 329-352.
Aaronson D, French E. 2009. The effect of progressive taxation on labor supply with hours and wages are jointly determined. Journal of Human Resources 44: 386-408.
Abowd J, Kramarz F, Margolis D. 1999. High wage workers and high wage firms. Econometrica, 67: 251-334.
Andrews DWK. 1988a. Chi-square diagnostic tests for econometric models: introduction and applications. Journal of Econometrics 37: 135-156.
Andrews DWK. 1988b. Chi-square diagnostic tests for econometric models: theory. Econometrica 56: 1419-1453.
Borjas G. 1980. The relationship between wages and weekly hours of work: the role of the division bias. Journal of Human Resources 15: 409-423.
Biddle J, Zarkin G. 1989. Choice among wage-hours packages: an empirical investigation of male labor supply. Journal of Labor Economics 7: 415-437.
Bloemen H. 2008. Job search, hours restrictions, and desired hours of work. Journal of Labor Economics 26: 137-179.
Bloemen H, Kapteyn A. 2008. The estimation of utility consistent labor supply models by means of simulated scores. Journal of Applied Econometrics 23: 395-422.
Blundell R, MaCurdy T, Meghir C. 2007. Labor supply models: unobserved heterogeneity, nonparticipation and dynamics. In Handbook of Econometrics, Vol. 6A, Heckman JJ, Leamer EE (eds). Elsevier: Amsterdam; 4667-4775.
Dagsvik JK. 1994. Discrete and continuous choice, max-stable processes and independence from irrelevant attributes. Econometrica 62: 1179-1205.
Dagsvik JK. 2013. Justification of functional form assumptions in structural models: a correction. Theory and Decision 75: 79-83.
Dagsvik JK, Jia Z. 2006. Labor supply as a choice among latent job opportunities: a practical empirical approach. Discussion paper 481, Research Department, Statistics Norway, Oslo.
Dagsvik JK, Røine Hoff S. 2011. Justification of functional form assumptions in structural models: applications and testing of qualitative measurement axioms. Theory and Decision 70: 215-254.
Dagsvik JK, Strøm S. 1997. A framework for labor supply analysis in the presence of complicated budget restrictions and qualitative opportunity aspects. Memorandum 22, Department of Economics, University of Oslo.
Dagsvik JK, Strøm S. 2004. Sectoral labor supply, choice restrictions and functional form. Discussion paper 388, Statistics Norway, Oslo.
Dagsvik JK, Strøm S. 2006. Sectoral labor supply, choice restrictions and functional form. Journal of Applied Econometrics 21: 803-826.
Dagsvik JK, Jia Z, Orsini K, van Camp G. 2011. Subsidies on low-skilled workers' social security contributions: the case of Belgium. Empirical Economics 40: 779-806.
Dagsvik JK, Jia Z, Kornstad T, Thoresen T. 2014. Theoretical and practical arguments for modeling labor supply as a choice among latent jobs. Journal of Economic Surveys 28: 134-151.
Di Tommaso ML, Sæther EM, Strøm S. 2009. Nurses wanted. Journal of Health Economics 28: 748-757.

Farzin YH. 2009. The effect of non-pecuniary motivations on labor supply. Quarterly Review of Economics and Finance 49: 1236-1259.
Hausman JA, Abrevaya J, Scott-Morton FM. 1998. Misclassification of the dependent variable in a discreteresponse setting. Journal of Econometrics 87: 239-269.
Kornstad T, Thoresen T. 2007. A discrete choice model for labor supply and child care. Journal of Population Economics 20: 781-803.
Lancaster K. 1966. A new approach to consumer demand. Journal of Political Economy 74: 132-157.
Lancaster K. 1971. Consumer Demand: A New Approach. Columbia University Press: New York.
Lundberg S. 1985. Tied wage-hours offers and the endogeneity of wages. Review of Economics and Statistics 83: 638-646.
Manning A, Petrongolo B. 2008. The part-time pay penalty for women in Britain. Economic Journal 118: F28-F51.
McFadden D. 1973. Conditional logit analysis of qualitative choice behavior. In Frontiers in Econometrics, Zarmebka P (ed.). Academic Press: New York; 105-142.
Moffitt R. 1984. The estimation of a joint wage-hours labor supply model. Journal of Labor Economics 2: 550-566.
Mortensen DT. 2003. Wage Dispersion: Why are Similar Workers Paid Differently? MIT Press: Cambridge, MA.
Sattinger M. 1993. Assignment models of the distribution of earnings. Journal of Economic Literature 31: 831-880.
Sattinger M. 1995. Search and the efficient assignment of workers to jobs. International Economic Review 36: 283-302.
Train K. 2009. Discrete Choice Methods with Simulation. Cambridge University Press: New York.
van Ophem H, Hartog J, Vijverberg W. 1993. Job complexities and wages. International Economic Review 34: 853-873.
van Soest A. 1995. Structural models of family labor supply: a discrete choice approach. Journal of Human Resources 30: 63-88.
van Soest A, Das M. 2001. Family labor supply and proposed tax reforms in the Netherlands. De Economist 149: 191-218.
van Soest A, Das M, Gong X. 2002. A structural labour supply model with flexible preferences. Journal of Econometrics 107: 345-375.
Wolf E. 2002. Lower wage rates for fewer hours? A simultaneous wage-hours model for Germany. Labour Economics 9: 643-663.

## APPENDIX A

Proof of Theorem 2. Let $F_{k}^{\prime}(x)$ denote the partial derivative with respect to the $k$ th component of $x$ of a function $F(x)$. By Assumption 3, $f(u, I)$ is invertible in $u$ for a given $I$ except at a finite number of points. Thus $\kappa(C, I)$ determined by $C=f(\kappa(C, I), I)$ is a well-defined function and equals the wage income which is needed to achieve disposable income $C$ at given non-labor income level $I$. From equation ( $2 \mathrm{a}, \mathrm{b}$ ) we have that

$$
\begin{equation*}
\frac{v(f(h w, I), h) \theta g(h, w)}{v(f(0, I), 0)}=\frac{\varphi(h, w \mid I)}{\varphi(0,0 \mid I)} \tag{A.1}
\end{equation*}
$$

where $\varphi(h, w \mid I)$ is the joint density of observed hours and wage rate. By taking the logarithm transformation of equation (A.1) and differentiating with respect to $I$ we obtain

$$
\begin{equation*}
\frac{v_{1}^{\prime}(f(h w, I), h) f_{2}^{\prime}(h w, I)}{v(f(h w, I), h)}-\frac{v_{1}^{\prime}(f(0, I), 0) f_{2}^{\prime}(0, I)}{v(f(0, I), 0)}=\frac{\varphi_{3}^{\prime}(h, w \mid I)}{\varphi(h, w \mid I)}-\frac{\varphi_{3}^{\prime}(0,0 \mid I)}{\varphi(0,0 \mid I)} \tag{A.2}
\end{equation*}
$$

By inserting $w h=\kappa(C, I)$ into equation (A.2) and rearranging, (A.2) becomes

$$
\begin{equation*}
\frac{v_{1}^{\prime}(C, h)}{v(C, h)}=\frac{\partial \log v(C, h)}{\partial C}=\frac{r(I)}{f_{2}^{\prime}(\kappa(C, I), I)}+\Gamma(C, h, I) \tag{A.3}
\end{equation*}
$$

where

$$
\begin{equation*}
\Gamma(C, h, I)=\left(\frac{\varphi_{3}^{\prime}(h, \kappa(C, I) / h \mid I)}{\varphi(h, \kappa(C, I) / h \mid I)}-\frac{\varphi_{3}^{\prime}(0,0 \mid I)}{\varphi(0,0 \mid I)}\right) \frac{1}{f_{2}^{\prime}(\kappa(C, I), I)} \tag{A.4}
\end{equation*}
$$

and

$$
\begin{equation*}
r(I)=v_{1}^{\prime}(f(0, I), 0) f_{2}^{\prime}(0, I) / v(f(0, I), 0) . \tag{A.5}
\end{equation*}
$$

By integrating equation (A.3) with respect to consumption we get that

$$
\begin{equation*}
\log v(C, h)-\log v\left(C_{0}, h\right)=r(I) \int_{C_{0}}^{C} \frac{\mathrm{~d} z}{f_{2}^{\prime}(\kappa(z, I), I)}+\int_{C_{0}}^{C} \Gamma(z, h, I) \mathrm{d} z \tag{A.6}
\end{equation*}
$$

Let
$\log \zeta(C, I)=\int_{C_{0}}^{C} \frac{\mathrm{~d} z}{f_{2}^{\prime}(\kappa(z, I), I)}$ and $\log \lambda^{*}(C, h, I)=\int_{C_{0}}^{C} \Gamma(z, h, I) \mathrm{d} z$
With this notation we obtain from equation (A.6) that

$$
\begin{equation*}
v(C, h)=\zeta(C, I)^{r(I)} \lambda^{*}(C, h, I) v\left(C_{0}, h\right) \tag{A.7}
\end{equation*}
$$

for $h>0$. We note that the left-hand side of equation (A.7) does not depend on $I$. We can therefore keep $I$ fixed and equal to any given $I^{*}$ (say) in the expression on the right-hand side. Since $f(u, I)$ is known and $\Gamma(C, h, I)$ is identified due to (A.4), both $\zeta\left(C, I^{*}\right)$ and $\lambda^{*}\left(C, h, I^{*}\right)$ are identified for positive $h$. However, even when $I^{*}$ has been fixed, equation (A.5) cannot be used to determine $r\left(I^{*}\right)$ since $v(C, 0)$ is not identified.
When $I^{*}$ has been fixed, we can suppress $I^{*}$ in the notation and write $\zeta(C)=\zeta\left(C, I^{*}\right), \lambda^{*}(C, h)=\lambda^{*}$ $\left(C, h, I^{*}\right), r=r\left(I^{*}\right)$ and $v\left(C_{0}, h\right)=\delta(h)$ in equation (A.7), which completes the proof.

Proof of Theorem 3. By Assumption 4, $g_{2}(w \mid h)=g_{2}(w)$. Note next that the relation in (A.7) does not necessesarily imply that $\mathrm{r}(\mathrm{I})$ is independent of I. From equations (2a) and (A.7) it then follows that for $I>0, h \neq h^{*} \in D$

$$
\frac{\varphi(h, w \mid I)}{\varphi\left(h^{*}, w \mid I\right)}=\frac{v(f(h w, I), h) g_{1}(h)}{v\left(f\left(h^{*} w, I\right), h\right) g_{1}\left(h^{*}\right)}=\frac{\zeta(f(h w, I), I)^{r(I)} \lambda^{*}(f(h w, I), h, I) \delta(h) g_{1}(h)}{\zeta\left(f\left(h^{*} w, I\right), I\right)^{r(I)} \lambda^{*}\left(f\left(h^{*} w, I\right), h^{*}, I\right) \delta\left(h^{*}\right) g_{1}\left(h^{*}\right)}
$$

which implies that for $w_{1} \neq w_{2}$

$$
\begin{align*}
& \log \left(\frac{\varphi\left(h, w_{1} \mid I\right) \lambda^{*}\left(f\left(h^{*} w_{1}, I\right), h^{*}, I\right)}{\varphi\left(h^{*}, w_{1} \mid I\right) \lambda^{*}\left(f\left(h w_{1}, I\right), h, I\right)} \cdot \frac{\varphi\left(h^{*}, w_{2} \mid I\right) \lambda^{*}\left(f\left(h w_{2}, I\right), h, I\right)}{\varphi\left(h, w_{2} \mid I\right) \lambda^{*}\left(f\left(h^{*} w_{2}, I\right), h^{*}, I\right)}\right)  \tag{A.8}\\
& \quad=r(I) \cdot \log \left(\frac{\zeta\left(f\left(h w_{1}, I\right), I\right)}{\zeta\left(f\left(h^{*} w_{1}, I\right), I\right)} \cdot \frac{\zeta\left(f\left(h^{*} w_{2}, I\right), I\right)}{\zeta\left(f\left(h w_{2}, I\right), I\right)}\right)
\end{align*}
$$

From equation (A.8) we see that $r(I)$ is identified. Thus $\lambda(C, h)$ defined by $\lambda(C, h)=\zeta(C, I)^{r(I)} \lambda^{*}(C, h, I)$ is identified and equation (A.7) implies that $\lambda(C, h)$ does not depend on $I$. Hence $v(C, h)=\lambda(C, h) \delta(h)$ for positive $h$. We thus realize that $v(C, h)$ is identified for positive $h$ up to a multiplicative term $\delta(h)$ that is solely a function of hours of work. Consider next the case with $h=0$. From equation (A.5) we have

$$
\begin{equation*}
\frac{\partial \log v(f(0, I), 0)}{\partial I}=r(I) \tag{A.9}
\end{equation*}
$$

The right-hand side of equation (A.9) is known, since $r(I)$ is identified by equation (A.8). By integrating equation (A.9) with respect to $I$ we realize that $\log v(f(0, I), 0)$ is determined for each positive $I$ up to an additive constant. But this means that $v(C, 0)$ is identified up to a multiplicative constant.

In the case where the distribution of offered wage rates is not degenerate we get from equation (2a) and the results above

$$
\begin{equation*}
g_{2}(w)=g_{2}\left(w^{*}\right) \frac{\varphi(h, w \mid I) / \lambda(f(h w, I), h)}{\varphi\left(h, w^{*} \mid I\right) / \lambda\left(f\left(h w^{*}, I\right), h\right)} \tag{A.10}
\end{equation*}
$$

By integrating equation (A.10) (or summing in the case with discrete offered wage rates) with respect to positive $w$ it follows (because the left-hand side is a p.d.f. function) that

$$
1=\frac{g_{2}\left(w^{*}\right)}{\varphi\left(h, w^{*} \mid I\right) / \lambda\left(f\left(h w^{*}, I\right), h\right)} \int_{0}^{\infty}(\varphi(h, y \mid I) / \lambda(f(h y, I), h)) \mathrm{d} y
$$

which together with (A.10) yields

$$
\begin{equation*}
g_{2}(w)=\frac{\varphi(h, w \mid I) / \lambda(f(h w, I), h)}{\int_{0}^{\infty}(\varphi(h, y \mid I) / \lambda(f(h y, I), h)) d y} \tag{A.11}
\end{equation*}
$$

which shows that $g_{2}(w)$ is identified. The proof is thus complete.

## APPENDIX B

## Simulation of Changes in the Opportunity Distribution of Offered Hours of Work

Recall that in our framework $g_{1 F}(h)$ represents the proportion of jobs with hours of work, $h$, that are available to the wife, whereas the parameter $\theta_{F}$ is a measure of job availability to the wife. Note that there are two peaks in the estimated offered hour distribution, namely when $h=h_{p}$ (part time) and $h=h_{f}$ (full-time). The proposed reform can be interpreted as a change in the opportunity distribution of hours for women, obtained by removing the part-time peak and increasing the fulltime peak (since the part-time jobs are replaced by full-time ones) while keeping the job availability unchanged (i.e. $\theta_{F}$ is unchanged). Denote the new opportunity distribution by $g_{1 F}^{*}(h)$. After the part-time peak has been removed, the offered hours are uniformly distributed apart from a peak at the full-time interval. Since there are five intervals for which the new opportunity density is constant, we must have that

$$
g_{1 F}^{*}(h)=\left(1-g_{1 F}\left(h_{p}\right)-g_{1 F}\left(h_{f}\right)\right) / 5
$$

for $h \neq h_{f}$. In addition, the sum of jobs in part-time and full-time intervals is the same before and after reform:

$$
\theta_{F} g_{1 F}\left(h_{p}\right)+\theta_{F} g_{1 F}\left(h_{F}\right)=\theta_{F} g_{1 F}^{*}\left(h_{p}\right)+\theta_{F} g_{1 F}^{*}\left(h_{F}\right)
$$

From these two equations it follows that

$$
g_{1 F}^{*}\left(h_{f}\right)=\left(6 g_{1 F}\left(h_{p}\right)+6 g_{1 F}\left(h_{f}\right)-1\right) / 5
$$

One can apply the model to simulate the corresponding realized labor supply distribution by replacing $g_{1 F}(h)$ with $g_{1 F}^{*}(h)$.


[^0]:    * Correspondence to: Zhiyang Jia, Research Department, Statistics Norway, Oslo, Norway. E-mail: zhiyang.jia@ssb.no
    ${ }^{1}$ Van Soest (1995) and others have proposed to introduce suitable dummies in the discrete labor supply model specification to improve the fit. However, this practice implies a non-structural model unless one interprets these dummies as part of the preference representation. In this case it means that the agent has stronger preferences for working particular hours (such as part-time and full-time hours) relative to other hours.

[^1]:    ${ }^{2}$ Dagsvik (1994) uses an equivalent additive utility specification, in which case the intensity that corresponds to $\theta \varepsilon^{-2}$ is equal to $\theta e^{-\varepsilon}$.
    ${ }^{3}$ For simplicity we apply the terminology of joint density function, although the marginal distribution of realized hours of work is discrete.

[^2]:    ${ }^{4}$ See for example, Bloemen (2008) and references therein for some recent studies using desired hours of work to identify preferences.

[^3]:    ${ }^{5}$ See also the errata in Dagsvik (2013).

[^4]:    ${ }^{6}$ The assumption is analogous in the case where $X$ is discrete, in which case the integral with respect to $X$ is replaced by a sum.

[^5]:    ${ }^{7}$ This error is sometimes called 'division bias' and is a rather common problem in many typically available datasets; see, for example, the discussion in Borjas (1980).
    ${ }^{8}$ One should, in principle, interpret the estimated error terms in the Mincer type wage equations as the sum of inter- and intraindividual effects, $\xi(z)+\eta$. Although it seems technically possible to separate the contribution of these two terms, we believe that this identification hinges too much on the mathematical structure of the model and is therefore not theoretically sound, given the available information.

[^6]:    ${ }^{10}$ Around $\$ 30$ per hour at 1997 exchange rates.

