Discussion Papers
Statistics Norway
Research department

No. 857
March 2017

Statistics Norway

Statistisk sentralbyrå

John K. Dagsvik and Steinar Strøm

Labor supply analysis with non-convex Budget sets without the Hausman approach

Discussion Papers No. 857, March 2017
Statistics Norway, Research Department

John K. Dagsvik and Steinar Strøm

Labor supply analysis with non-convex Budget sets without the Hausman approach

Abstract:
When the budget set is non-convex the application of the Hausman approach to estimate labor supply functions will in general be cumbersome because labor supply no longer depends solely on marginal criteria (first order conditions). In this paper we demonstrate that the conventional continuous labor supply model (including corner solution for non-participation) with non-convex budget sets in some cases can be estimated using only first order conditions provided the budget curve is continuously differentiable and the utility function belongs to a particular class. We subsequently discuss how the model can be specified econometrically. Finally, we discuss the application of the model to simulate the effect of counterfactual reforms.

Keywords: Labor supply, non-convex budget sets, marginal criteria

JEL classification: C51, J22

Acknowledgements: We are grateful for comments by Terje Skjerpen

Address: John K. Dagsvik, Statistics Norway, Research Department. E-mail: john.dagsvik@ssb.no

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

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

ISSN 1892-753X (electronic)

Sammendrag

Modellering av arbeidstilbud ved å ta utgangspunkt i den konvensjonelle lærebok-tilnærmingen er komplisert i tilfellet med ikke-convekse budsjett mengder fordi tilbudet ikke lengre avhenger kun av marginale kriterier (første-ordens betingelser). I denne artikkelen viser vi at den konvensjonelle kontinuerlige arbeidstilbudsmodellen (inkludert hjørneløsning for ikke-arbeid-alternativet) med ikke-konvekse budsjettmenger i kan estimeres ved kun å benytte første-ordens betingelser så fremt budsjettkurven er kontinuerlig deriverbar og nyttefunksjonen tilhører en bestemt klasse. Deretter diskuterer vi hvordan modellen kan spesifiseres økonometrisk. Endelig drøfter vi noen aspekter ved anvendelse av modellen til å simulere kontrafaktiske reformer.

1. Introduction

The traditional models of labor supply are based on a version of the theory of consumer behavior with two goods, namely consumption (disposable income) and leisure. Whereas this theory seems straightly forward to apply empirically in the context of consumer demand, it is much less so when it comes to analyses of labor supply. There are several reasons for this. One reason is that due to the economic budget constraint the budget set will be kinked and non-convex and accordingly the usual marginal calculus does not apply. The so-called Hausman approach, initiated by Burtless and Hausman (1978), Hausman (1979) and Hausman (1985a, b), enables the researcher to account for kinks and non-convexity of the budget constraint. Unfortunately, the Hausman approach has proved to be very hard to apply in practice, see for example Bloemen and Kapteyn (2008).[1] A major reason for the difficulty inherent with empirical applications of the Hausman approach is that the usual first order conditions are no longer sufficient for determining optimal hours of work. In general, with non-convex budget sets, one needs global criteria for determining optimal hours of work.

The purpose of this paper is to demonstrate that with the traditional (continuous) labor supply modelling approach (allowing for non-participation) it suffice to use first order conditions to determine labor supply provided the tax function is smooth and the marginal tax rate at zero hours equals zero and the utility function belongs to a particular class. Subsequently, we discuss a possible econometric specification and how one can estimate the model without solving explicitly for the labor supply function.

Recently, the so-called discrete choice labor supply models have become popular, see van Soest (1995). The discrete choice models can readily handle any nonlinear budget constraints and rather general utility specifications. The discrete choice labor supply models can approximate a continuous model by increasing the number of feasible discrete hours of work. The essential difference between traditional continuous and discrete labor supply models is the stochastic specification of the stochastic terms that enter the utility function. Still, we believe it is of interest to discuss the conventional approach as an alternative for modeling labor supply. It is an empirical question whether or not a model based on continuous choice (with restricted specification on the stochastic terms) is better than a discrete choice model based on the discrete choice framework, à la van Soest (1995).[2]

---

[1] See also Heim (2009) who demonstrate how a simulation procedure can be applied to estimate labor supply models with kinks provided the budget set is convex.

[2] A weakness with both the traditional discrete and continuous choice models is their inability to handle restrictions on hours of work. Dagsvik et al. (1988) proposed a model based on the notion of latent job choice, which was further developed by Dagsvik and Strøm (2006) and Dagsvik and Jia (2016), which can accommodate

In the next section the model is analyzed in the case where the budget set is non-convex. In Section 3 we specify how the model can be estimated. Section 4 discusses how the model can be used to carry out different simulations of counterfactual reforms.

2. The model

In this section we discuss our particular approach based on a marginal criterion approach. Let $U(C,h)$ denote the agent's utility in consumption C (disposable income) and annual hours of work, where $0 \le h \le M$ and M is the maximum hours of work the agent can work. Let T denote the tax function and let

(2.1) $$C(h) = Wh + I - T(Wh, I)$$

where W is the agent's wage rate and I his non-labor income.

Assumption 1

The utility function has the structure

(2.2) $$U(C,h) = V_1(C) + \kappa V_2(h) + r\kappa V_1(C)V_2(h)$$

where κ is a positive random variable, r is a constant, $V_1(C)$ is a continuously differentiable, strictly increasing and concave function of C and $V_2(h)$ a continuously differentiable, strictly decreasing and concave function of h.

Note that the assumption that $\kappa > 0$ represents no restriction because it is the sign of $\kappa V_2(h)$ that matters in the model. Unfortunately, Assumption 1 is, however, not sufficient to ensure quasi-concavity of the utility function. We have the following result.

Proposition 1

Under Assumption 1 a necessary and sufficient condition for the utility function to be strictly increasing in C and strictly decreasing in h is that

(2.3) $$1 + r\kappa V_2(h) > 0 \text{ and } 1 + rV_1(C) > 0.$$

If (2.3) holds the utility function in (2.2) is quasi-concave when

(2.4) $$r\kappa \ge \frac{(V_1'(C))^2 V_2''(h) + \kappa (V_2'(h))^2 V_1''(h)}{2(V_1'(C)V_2'(h))^2 - V_2''(h)V_2(h)(V_1'(C))^2 - V_1''(C)V_1(C)(V_2'(h))^2}.$$

---

restrictions on hours of work and job opportunities. The framework of Dagsvik and Strøm (2006) is therefore suitable to apply when choices are constrained.

The proof of Proposition 1 is given in the Appendix. Note that the right hand side of (2.4) is always negative. Thus, the next result follows.

Corollary 1

If (2.3) holds the utility in (2.2) is always quasi-concave when $r \geq 0$.

Assumption 1 is fairly general and is always satisfied by utility functions that are separable in consumption and hours of work provided they are increasing and concave in consumption and decreasing and concave in hours. It is however, not satisfied by the utility function typically applied in the Hausman approach, which yields a linear labor supply function. It is not satisfied by the general quadratic utility function either (Heim, 2009) apart from the case where utility is separable in hours and consumption.

An interesting special case of utility functions within the class defined by (2.2) is obtained when $V_1(C)$ and $V_2(h)$ are given by the Box-Cox functions

$$(2.5) \qquad V_1(C) = \frac{C^{\alpha_1} - 1}{\alpha_1} \quad \text{and} \quad V_2(h) = \frac{(M-h)^{\alpha_2} - 1}{\alpha_2}.$$

A justification of the utility structure given by (2.2) and (2.5) based on particular invariance properties has been given by Dagsvik and Røine Hoff (2011). When (2.5) holds the condition in (2.4) reduces to

$$r\kappa \geq -\frac{\kappa(1-\alpha_1)C^{-\alpha_1} + (1-\alpha_2)(M-h)^{-\alpha_2}}{\kappa(1-\alpha_1)C^{-\alpha_1}V_1(C) + (1-\alpha_2)(M-h)^{-\alpha_2}V_2(h) + 2}.$$

Let m be the number of tax segments and let h_j denote the hours of work that corresponds to the j-th kink point of the budget curve $xW + I - T(xW, I),\ x \in [0, M].$ For $h \in [h_{j-1}, h_j]$ let

$$B_j(h) = \{x \in [h_{j-1}, h_j] : V_2(x)(1 + rV_1(C(x))) \geq V_2(h)(1 + rV_1(C(h)))\}$$

and

$$\bar{B}_j(h) = \{x \in [h_{j-1}, h_j] : V_2(x)(1 + rV_1(C(x))) < V_2(h)(1 + rV_1(C(h)))\}$$

for $j > 0$ and $h_0 = 0,$ and define

$$(2.6) \qquad g(h) = \log\left(\frac{-V_1'(C(h))C'(h)}{V_2'(h)(1 + rV_1(C(h))) + rC'(h)V_1'(C(h))V_2(h)}\right),$$

$$(2.7a) \qquad \psi_{1j}(h) = \min_{x \in B_j(h)} \log\left(\frac{V_1(C(h)) - V_1(C(x))}{V_2(x)(1 + rV_1(C(x))) - V_2(h)(1 + rV_1(C(h)))}\right)$$

and

(2.7b) $$\psi_{2j}(h) = \max_{x \in \bar{B}_j(h)} \log\left(\frac{V_1(C(h)) - V_1(C(x))}{V_2(x)(1+rV_1(C(x))) - V_2(h)(1+rV_1(C(h)))}\right).$$

Note that when h is positive $g(h)$ depends on W through the wage rate which enters in (2.1).[3]

In the case with non-convex budget sets there may be multiple tangencies between the marginal rate of substitution and the budget curve.

Assumption 2

The tax function is piecewise linear and the marginal tax rate is equal to zero at zero hours of work.

The property that the marginal tax rate is zero for hours of work equal to zero is a feature that is shared with practically all tax systems. In the following let $\tilde{h}$ denote the chosen hours of work. We have:

Theorem 1

Assume that Assumptions 1 and 2 hold. Then hours of work $\tilde{h} \in (h_{j-1}, h_j)$ *if* $\log \kappa = g(\tilde{h})$ *and*

$$\max_{k \neq j} \psi_{2k}(\tilde{h}) \le g(\tilde{h}) \le \min_{k \neq j} \psi_{1k}(\tilde{h}).$$

Hours of work is located at kink point $h_j, j > 0,$ *if*

$$\max(g(h_j -), \max_{k \notin \{j-1,j,j+1\}} \psi_{2k}(h_j)) < \log \kappa < \min(g(h_j +), \min_{k \notin \{j-1,j,j+1\}} \psi_{1k}(h_j)).$$

Hours of works equals zero if $\log \kappa \le g(0)$.

Proof:

The first order condition implies that optimal hours of work subject to the budget constrain (2.1) is given by

$$C'(\tilde{h}) \le -\frac{U_2'(C(\tilde{h}), \tilde{h})}{U_1'(C(\tilde{h}), \tilde{h})}$$

where equality holds when $\tilde{h}$ is positive and different from the kink points. Hence, in the case of interior solution $\tilde{h} \in (h_{j-1}, h_j)$ we must have that

[3] Note that the denominator in the expression on the right hand side of (2.6) is decreasing in h (unless for extreme tax systems). Hence, if the denominator is negative for interior solutions to occur (which it must to allow for interior solutions) it will therefore be negative for $h = 0$.

$$C'(\tilde{h}) = -\frac{U_2'(C(\tilde{h}),\tilde{h})}{U_1'(C(\tilde{h}),\tilde{h})}.$$

It follows readily that under Assumption 1 the first order conditions above yield $g(\tilde{h}) = \log\kappa$. In order for hours of work to be the optimal choice within (h_{j-1}, h_j) it must be the case that

$U(C(\tilde{h}),\tilde{h}) \geq U(C(h),h)$ for all $h \notin [h_{j-1}, h_j]$. Under Assumption 1 the inequality above becomes:

$$V_1(C(\tilde{h})) + \kappa V_2(\tilde{h}) + r\kappa V_1(C(\tilde{h}))V_2(\tilde{h}) \geq V_1(C(h)) + \kappa V_2(h) + r\kappa V_1(C(h))V_2(h)$$

for $h \in [0, M] \setminus [h_{j-1}, h_j]$. By (2.7 a,b) the latter inequality is equivalent to

(2.8) $$\max_{k \neq j} \psi_{2k}(\tilde{h}) \leq \log\kappa \leq \min_{k \neq j} \psi_{1k}(\tilde{h}).$$

Since $g(\tilde{h}) = \log\kappa$ in the case of positive *h* different from the kink points we obtain from (2.8) that

$$\max_{k \neq j} \psi_{2k}(\tilde{h}) \leq g(\tilde{h}) \leq \min_{k \neq j} \psi_{1k}(\tilde{h})$$

which proves the first part of the theorem.

Consider the case where $\tilde{h}$ is located at the kink point h_j. For this to happen it must be the case that

$$C'(h_j+) < -\frac{U_2'(C(h_j),h_j)}{U_1'(C(h_j),h_j)} < C'(h_j-)$$

which is equivalent to[4]

(2.9) $$g(h_j-) < \log\kappa < g(h_j+).$$

In addition, one must have that

$$V_1(C(h_j)) + \kappa V_2(h_j) + r\kappa V_1(C(h_j))V_2(h_j) \geq V_1(C(h)) + \kappa V_2(h) + r\kappa V_1(C(h))V_2(h)$$

for $h \in [0, M] \setminus [h_{j-2}, h_{j+1}]$. The latter inequality is equivalent to

(2.10) $$\max_{k \notin \{j-1,j,j+1\}} \psi_{2k}(h_j) < \log\kappa < \min_{k \notin \{j-1,j,j+1\}} \psi_{1k}(h_j).$$

If we combine (2.9) and (2.10) we obtain the second part of the theorem.

Consider finally the case where the optimal decision is not to work. By Assumption 2 the tax function is such that the marginal tax rate at zero hours is zero and thus the marginal rate of substitution at zero hours of work is greater than the marginal rate of substitution at any tax segment. It follows that if $g(0) \geq \log\kappa$ the optimal decision is not to work. Hence, the result of the theorem follows.

Q.E.D.

---

[4] As usual, the notation $h+$ means $\lim_{h \to h_j} f(h) = f(h_j+)$. The definition of h_j- is similar.

We next consider choice probabilities. Let $f(x)$ and $F(x)$ be the conditional p.d.f. and conditional c.d.f. of $\log\kappa$ given the wage rate. Also let $1\{A\}$ denote the indicator function that is equal to 1 if the event A occurs and zero otherwise.

Theorem 2

Suppose that Assumptions 1 and 2 hold. Then the conditional probability density of hours of work, given the wage rate, is equal to

$$\lambda(h) = f(g(h))\,|\,g'(h)\,|$$

for $h > 0$, *provided h is not a kink point. The conditional probability of* $\tilde{h}$ *being located at kink point* h_j, *given the wage rate is equal to*

$$P(\tilde{h} = h_j) = F\Big(\min(g(h_j+), \min_{k\notin\{j-1,j,j+1\}} \psi_{1k}(h_j))\Big) - F\Big(\max(g(h_j-), \max_{k\notin\{j-1,j,j+1\}} \psi_{2k}(h_j))\Big).$$

The conditional probability of $\tilde{h} = 0$, *given the wage rate, is equal to*

$$P(\tilde{h} = 0) = F(g(0)).$$

Proof: Consider first the case where $\tilde{h} \in (h_{j-1}, h_j), j > 0.$ Note that the function $g(h),$ as a function of h, is invertible in $(h_{j-1}, h_j).$ To determine which of the segments is the one that corresponds to the observed hours of work, we apply Theorem 1. From Theorem 1 it follows, with a small Δh and $h, h + \Delta h \in (h_{j-1}, h_j),$ that

$$P(\tilde{h} \in (h, h+\Delta h)) = f(g(h))\,|\,g'(h)\,|\,1\{\max_{k\neq j} \psi_{2k}(h) \le g(h) \le \min_{k\neq j} \psi_{1k}(h)\}\Delta h + o(\Delta h).$$

Note next that $1\{\max_{k\neq j} \psi_{2k}(h) \le g(h) \le \min_{k\neq j} \psi_{1k}(h)\}$ is equal to 1 or 0 because $g(h)$ and $\psi_{kr}(h), k = 1,2, r = 1,2,...,m,$ are deterministic functions. If $1\{\max_{k\neq j} \psi_{2k}(h) \le g(h) \le \min_{k\neq j} \psi_{1k}(h)\} = 0$ then h cannot be the optimal choice and therefore we conclude that $1\{\max_{k\neq j} \psi_{2k}(h) \le g(h) \le \min_{k\neq j} \psi_{1k}(h)\} = 1$ in which case this factor does not play any role in the choice probability above. Hence we obtain that

$$P(\tilde{h} \in (h, h+\Delta h)) = f(g(h))\,|\,g'(h)\,|\,\Delta h + o(\Delta h).$$

Thus, it follows that the corresponding p.d.f. is given by $f(g(h))\,|\,g'(h)\,|.$

Consider next the probability that the $\tilde{h}$ is located at kink point $h_j, j > 0.$ From Theorem 1 it follows that

$$P(\tilde{h}=h_j)=P\Big(\max(g(h_j-),\max_{k\notin\{j-1,j,j+1\}}\psi_{2k}(h_j))<\log\kappa<\min(g(h_j+),\min_{k\notin\{j-1,j,j+1\}}\psi_{1k}(h_j))\Big)$$

$$=F\Big(\min(g(h_j+),\min_{k\notin\{j-1,j,j+1\}}\psi_{1k}(h_j))\Big)-F\Big(\max(g(h_j-),\max_{k\notin\{j-1,j,j+1\}}\psi_{2k}(h_j))\Big).$$

Consider finally the case where $\tilde{h}=0.$ Then from Theorem 1 it follows that

$$P(\tilde{h}=0)=P(g(0)\geq\log\kappa)=F(g(0)).$$

Hence, the proof is complete.

Q.E.D.

Theorem 2 implies that in the presence of kinks in the budget curve the likelihood function becomes non-differentiable. However, if only observations on participation versus non-participation are used then the corresponding likelihood function becomes differentiable. This is an implication of the next corollary.

Corollary 2

Under Assumptions 1 and 2 the conditional probabilities of working and not working given the wage rate equal

$$P(\tilde{h}>0)=1-F(g(0)) \quad \textit{and} \quad P(\tilde{h}=0)=F(g(0)).$$

In principle one could apply the result of Theorem 2 to develop a full information maximum likelihood procedure. Unfortunately, there are several problems with such an endeavor. First, if observations on kink points are present the likelihood function will be non-differentiable on some subset of the parameter space. Although this problem might be overcome by recent developments in statistical theory (Drton, 2009), a more serious problem is perhaps the fact that the computation of the likelihood function involves the minimization and maximization of functions over a continuous interval.

The next result is an immediate implication of Theorem 2.

Corollary 3

Suppose that the tax function is continuously differentiable and that Assumptions 1 and 2 hold. Then the conditional probability density of hours of work, given the wage rate, is equal to

$$\lambda(h)=f(g(h))\,|\,g'(h)\,|$$

for $h>0,$ *provided h is not a kink point. The conditional probability of* $\tilde{h}=0$ *given the wage rate is equal to*

$$P(\tilde{h}=0)=F(g(0)).$$

Corollary 3 implies that the problem with multiple tangencies can easily be dealt with in the case with continuously differentiable budget curves.[5] This is due to the particular structure of the utility function asserted in Assumption 1. In empirical analyses one must extend the result of Theorem 2 to account for the fact that the wage rate is unobserved for those who do not work. In addition κ may be correlated with W due to the fact that individuals which are attractive on the labor market, and may therefore be offered high wage rates, may also have strong preferences for supplying labor. We shall discuss these aspects further in the next section.

Note that the result obtained in Theorems 1 and 2 do not depend critically on κ being the only source of randomness. Suppose for example that $g(h)=g(h,\mu)$ and $\psi_k(h)=\psi_k(h,\mu)$ where μ is a random vector that we for simplicity assume is independent of κ. Then, conditional on μ the p.d.f. of the optimal hours of work in case of an interior solution different from the kink points will, according to Theorem 2 be equal to $f(g(h,\mu))\,|\,g_1'(h,\mu)\,|$ where $g_1'(h,\mu)$ denotes the derivative with respect to *h*. Thus, the corresponding unconditional density will be $E(f(g(h,\mu))\,|\,g_1'(h,\mu)\,|)$ where the expectation is taken with respect to μ. Similarly, it follows from Theorem 2 that the probability of not working in this case becomes equal to $E(F(g(0,\mu)))$.

It should be noted that once the model has been estimated one can use the utility function directly when assessing the effect of counterfactual policies such as tax reforms, for example. This means that one does not need to employ the rather cumbersome first order conditions to predict the effect of counterfactual reforms. Moreover, the explicit form of the utility function allows for calculation of welfare in terms of compensating variations.

3. Econometric specification and maximum likelihood estimation with generalized Box-Cox utility function

In this section we consider econometric specification and maximum likelihood estimation when the budget curve is assumed to be continuously differentiable and the utility function is given by

$$U(C,h)=V_1(C-\overline{C})+\kappa V_2(h)$$

[5] Saez (2010) has analyzed bunching of taxpayers at kink points of the US income tax schedule. He has found clear evidence of bunching only around the first kink point for the self-employed but no bunching for wage earners. Saez suggests that tax evasion may be the reason for the bunching at the first kink point for the self-employed.

where $V_1(C)$ and $V_2(h)$ are given in (2.5) and

(3.1) $$\kappa = \exp(X_2 b + \varepsilon_2),$$

where $\bar{C}$ is subsistence consumption, ε_2 is a random error term and X_2 is a vector of individual characteristics that are assumed to affect the preference for leisure. In addition, we extend the analysis above by allowing for endogenous wage rates. Note that the utility function specified above belongs to the class given in (2.2) with $r = 0.$ The general case with $r \neq 0$ can be treated in a similar way.

In order to estimate the model we need to introduce a wage equation. We assume that the wage rate is given by the wage equation

(3.2) $$\log W = X_1\gamma + \varepsilon_1,$$

where X_1 is a vector of individual characteristics and γ is an unknown parameter vector and ε_1 is an error term. We assume further that $(\varepsilon_1, \varepsilon_2)$ are bivariate normally distributed error terms with zero mean. Because of this assumption we can write:

(3.3) $$\varepsilon_2 = \theta\varepsilon_1 + \varepsilon_3$$

where ε_3 is a zero mean normal random variable that is independent of ε_1, and θ is a constant. The covariance between the error term in the wage equation and the taste shifter is then given by $Cov(\varepsilon_1, \varepsilon_2) = \theta\sigma_1^2$, where $\sigma_1^2 = Var\varepsilon_1$. We shall use (3.3) in order to apply a version of the control function approach (see Wooldridge, 2010). From (2.6) it follows that

(3.4) $$g(\tilde{h}) = (\alpha_1 - 1)\log(C(\tilde{h}) - \bar{C}) + (1 - \alpha_2)\log(M - \tilde{h}) + \log(1 - T_1'(W\tilde{h}, I)) + \log W.$$

From Theorem 1 and (3.1) it follows that

(3.5) $$\varepsilon_2 = g(\tilde{h}) - X_2 b.$$

Let $X = (X_1, X_2)$ and

$$\begin{aligned}\xi(\tilde{h}, W, I, X) &= g(\tilde{h}) - \log W - X_2 b \\ &= (\alpha_1 - 1)\log(C(\tilde{h}) - \bar{C}) + (1 - \alpha_2)\log(M - \tilde{h}) + \log(1 - T_1'(W\tilde{h}, I)) - X_2 b.\end{aligned}$$

From (3.3) and (3.2) we obtain that

(3.6) $$\varepsilon_3 = g(\tilde{h}) - X_2 b - \theta(\log W - X_1\gamma) = \xi(\tilde{h}, W, I, X) + (1 - \theta)\log W + \theta X_1\gamma$$

It follows from (3.3) that the error term ε_3 is independent of *W*. Thus, econometrically, *W* appears in relation (3.6) as if it were exogenous. Similarly to Theorem 1 it follows from (3.6) that the p.d.f. of $\tilde{h}$ is given by

(3.7) $$\lambda(h \mid W, X, I) = \frac{1}{\sigma_3}\varphi\left(\frac{\xi(h, W, I, X) + (1 - \theta)\log W + \theta X_1\gamma}{\sigma_3}\right)\left|\xi_1'(h, W, I, X)\right|$$

for $h > 0$ where φ is the standard normal p.d.f. and[6]

$$g_1'(h) = \xi_1'(h, W, I, X) = \frac{(\alpha_1 - 1)W(1 - T_1'(Wh, I))}{C(\tilde{h}) - \overline{C}} + \frac{\alpha_2 - 1}{M - h} - \frac{WT_1''(Wh, I)}{1 - T_1'(Wh, I)}.$$

Moreover, if $\varepsilon_2 < g(0) - X_2 b$ the optimal decision is not to work. Using (3.2) and (3.3) the last inequality is equivalent to

(3.8) $$\varepsilon_3 + (\theta - 1)\varepsilon_1 < \xi(0, 0, I, X) + X_1\gamma$$

Hence, (3.8) implies that

(3.9) $$P(\tilde{h} = 0 \mid I, X) = \Phi\left(\frac{\xi(0, 0, I, X) + X_1\gamma}{\sqrt{\sigma_3^2 + (1 - \theta)^2 \sigma_1^2}}\right)$$

where Φ is the standard normal c.d.f. We note that in this case the marginal tax rate at zero hours of work is equal to zero, which is one of the requirements in Assumption 2. Let S_1 be the subsample of those who work and S_0 the subsample of those who do not work. From (3.7) and (3.9) it follows immediately that the log likelihood function for those individuals who work is given by

(3.10) $$L_1 = \prod_{i \in S_1}\left(\frac{1}{\sigma_3}\varphi\left(\frac{\xi(\tilde{h}_i, W_i, I_i, X_i) + (1 - \theta)\log W_i + \theta X_{1i}\gamma}{\sigma_3}\right) | \xi_1'(\tilde{h}_i, W_i, I_i, X_i) | \varphi\left(\frac{\log W_i - X_{1i}\gamma}{\sigma_1}\right)\frac{1}{\sigma_1}\right)$$

where the subscript *i* indexes the individuals. The likelihood function for those who do not work is given by

(3.11) $$L_0 = \prod_{i \in S_0}\Phi\left(\frac{\xi(0, 0, I_i, X_i) + X_{1i}\gamma}{\sqrt{\sigma_3^2 + (1 - \theta)^2 \sigma_1^2}}\right).$$

The total likelihood is thus the product of the likelihood functions in (3.10) and (3.11). We note that the likelihood functions in (3.10) and (3.11) are much simpler than the corresponding likelihood function that is implied by the Hausman approach (Bloemen and Kapteyn, 2008, and Heim, 2009).

4. Counterfactual reforms and labor supply responses

In this section we discuss how the model above can be applied for the purpose of simulating the effect of hypothetical interventions such as changes in the wage rate, non-labor income or tax rules. We believe there are two ways in which the model can be used.

[6] Note that the two first parts on the right hand side in the above equation is in fact the Slutsky elasticity.

The first type of simulation, and perhaps the most interesting one, is to use the model as a *pure labor supply model*. Recall that the role of the wage equation is analogous to an instrument variable relation which is only supposed to represent the distribution of the wage rates in the actual labor market which the individuals (in the sample) are facing, as well as the correlation between the error terms in the wage equation and the utility representation. Once the model has been estimated the wage equation no longer has any role to play in this case where the purpose is to simulate labor supply responses given a hypothetical intervention. However, since the utility function contains a stochastic part, representing unobserved heterogeneity in preferences, it is only possible to simulate labor supply responses in a distributional sense. Since the utility function has been estimated there is no need to use the marginal calculus of Theorem 2 to simulate counterfactual responses. The following simulation procedure is easy to apply: For individual *i* draw independent normal taste shifter ε_2 with standard deviation σ_2. Second, individual characteristics, wage rates, non-labor incomes and the tax system must be selected for each individual in the population. Then, for each individual the estimated utility function is maximized numerically, given the simulated random error term. This is done by first approximating the continuous set of possible hours of work $[0, M]$ by an associated finite set with a large number of points and subsequently determining labor supply as the value of hours of work that maximizes utility on the associated finite set of hours. However, if one is only interested in simulating labor market participation this can be done using a simple formula as we shall describe below.

A second way of conducting simulations from hypothetical intervention aims at simulating *realized* labor supply in the labor market. This approach rests on the assumption that, the structure of the wage equation holds, apart from the intercept, also in the new labor market equilibrium (including the joint distribution of $(\varepsilon_1, \varepsilon_2)$ after a change in the intercept of the wage equation, non-labor income or tax system has occurred. Recall, however, that the wage equation is not a fully structural equation because it is silent about how the wage rates respond to the policy of labor market unions, macro-economic variables such as import and export prices, interest rates, firm productivity and tax changes in rates. In principle one could use panel data or repeated cross-section data to test whether or not the parameters of the wage equation are constant over time. Furthermore, similarly to Heckman and Sedlaceck (1985), one could use macro time series data to estimate a structural specification of the intercept in the wage equation.

To illustrate the difference between the simulations of different counterfactual settings we shall consider simulation of labor market participation. Consider first the case where one wishes to predict parcitipation for a a population of women with characteristics X_2 facing wage rate *W*. Note first that when σ_1, σ_3 and θ have been estimated one can compute σ_2 by using the relation $\sigma_2^2 = \theta^2 \sigma_1^2 + \sigma_3^2$.

The latter equation follows from (3.3). From (3.4) and the condition $\varepsilon_2 < g(0) - X_2 b$ it follows that the participation probability is given by

$$P(\varepsilon_2 < g(0) - X_2 b) = \Phi\left(\frac{(\alpha_1 - 1)\log(C(0) - \bar{C}) + (1 - \alpha_2)\log M + \log W - X_2 b}{\sigma_2}\right).$$

The expression above predicts the fraction of women who wish to work given that they all have the same observable preference characteristics X_2 and given that the (hypothetical) wage rate is equal to W.

Next, consider the setting where one wishes to predict pure participation for a population of women with characertistics (X_1, X_2) where the women face different wage rates, as predicted by a wage equation which includes an error term that is normally distributed with hypothetical coefficient vector $\tilde{\gamma}$ associated with X_1 and variance $\tilde{\sigma}_1^2$ of the hypothetical error term. In this case it follows that the participation probability is given by

$$\Phi\left(\frac{(\alpha_1 - 1)\log(C(0) - \bar{C}) + (1 - \alpha_2)\log M + X_1\tilde{\gamma} - X_2 b}{\sqrt{\sigma_2^2 + \tilde{\sigma}_1^2}}\right).$$

The expression above is the fraction of women with characteristics (X_1, X_2) who would like to work given the distribution of wage rates described by the hypothetical wage equation. Note that here, the parameters $\tilde{\gamma}$ and $\tilde{\sigma}_1$ are not necessarily equal to the corresponding estimates obtained in empirical analyses. For example, it may be of interest to let $\tilde{\sigma}_1 = 0$ which means that the hypothetical wage rate for each woman is assigned without error.

Consider finally the case where one wishes to predict realized labor supply in a labor market that is similar to the one from which the data have been obtained. As mentioned above, in this scenario the maintained assumptions about the wage rates are as in the empirical situation. Note first that $\rho = \theta\sigma_1 / \sigma_2$ implying that $\sigma_3^2 + (1-\theta)^2\sigma_1^2 = \sigma_2^2 + \sigma_1^2 - 2\rho\sigma_1\sigma_2$. From (3.9) the fraction of women with characteristics (X_1, X_2) who will realize participation equals

$$\Phi\left(\frac{(\alpha_1 - 1)\log(C(0) - \bar{C}) + (1 - \alpha_2)\log M + X_1\gamma - X_2 b}{\sqrt{\sigma_2^2 + \sigma_1^2 - 2\rho\sigma_1\sigma_2}}\right).$$

In the latter case we have taken into account that the error term in the utility function and the wage rates are correlated (ceteris paribus) because individuals with high unobserved abilities and human capital are attractive on the labor market and consequently will receive high wage offers. Suppose for example that $\tilde{\gamma} = \gamma$ and $\tilde{\sigma}_1 = \sigma_1$. Then, if

$$(\alpha_1 - 1)\log(C(0) - \bar{C}) + (1 - \alpha_2)\log M + X_1\gamma - X_2 b > 0$$

the fraction of women that will realize participation is actually higher than the fraction of women who wishes to participate, despite the assumption that the distribution of wage rates in the two cases are equal. At first glance this may seem counterintuitive. The reaon is that the wage rates in the pure supply situation are assigned with errors which are uncorrelated with the error term in the reservation wage rate. In contrast, in the actual labor market

$$E(\log W \mid \varepsilon_2) = E\log W + E(\varepsilon_1 \mid \varepsilon_2) = E\log W + \rho\sigma_1 / \sigma_2$$

which shows that when controlling for unobserved abilities the average offered wage rate (under the same conditions as in the labor labor market that generated the data) is systematically higher than the average hypothetical wage rate in the pure labor supply scenario.

5. Conclusion

The conventional neoclassical labor supply model in the presence of tax systems that generate non-convex budget sets has proved difficult to estimate. In this paper we have demonstrated that it is possible, under rather general conditions, to estimate labor supply models using conventional techniques without applying the Hausman (or Heim) methodology. Our methodology does not require a closed form solution for the labor supply function.

Finally, we have discussed the application of the model to predict labor supply in different counterfactual scenarios.

References

Bloemen, H. and A. Kapteyn (2008): The Estimation of Utility Consistent Labor Supply Models by Means of Simulated Scores. *Journal of Applied Econometrics*, **23**, 395 – 422.

Burtless, G. and J. A. Hausman (1978): The Effect of Taxation on Labor Supply: Evaluating the Gary Negative Income Experiment. *Journal of Political Economy*, **86**, 1103–1130.

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

Dagsvik, J. K. and S. Røine Hoff (2011): Justification of Functional Form Assumptions in Structural Models: Applications and Testing of Qualitative Measurement Axioms. *Theory and Decision*, **70**, 215–254.

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 Discrete Choice among Latent Jobs: Unobserved Heterogeneity and Identification. *Journal of Applied Econometrics*, **31**, 487–506.

Drton, M. (2009): Likelihood Ratio Tests and Singularities. *Annals of Statistics*, **37**, 979–1012.

Hausman, J. A. (1979): The Econometrics of Labor Supply on Convex Budget Sets. *Economic Letters*, **3**, 171–174.

Hausman, J. A. (1985a): Taxes and Labor Supply, in *Handbook of Public Economics* Vol 1, (eds.) A. J. Auerbach and M. Feldstein, Elsevier Science Publisher B.V. (North–Holland), 213–264.

Hausman, J. A. (1985b): The Econometrics of Nonlinear Budget Sets. *Econometrica*, **53**, 1255–1282.

Heckman, J. J. and G. Sedlacek (1985): Heterogeneity, Aggregation and Market Wage Function: An Empirical Model of Self-Selection in the Labor Market. *Journal of Political Economy*, **93**, 1077–1125.

Heim, B. (2009): Structural Estimation of Family Labor Supply with Taxes: Estimating a Continuous Hours Model Using a Direct Utility Specification. *Journal of Human Resources*, **44**, 350–385.

Saez, E. (2010): Do Taxpayers Bunch at Kink Points? *American Economic Journal: Economic Policy*, **2**, 180–212.

Wooldridge, J. M. (2010): *Econometric Analysis of Cross Section and Panel Data*. Second edition. MIT Press, Cambridge, Massachusetts.

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

Appendix

Proof of Proposition 1:

Let $U(C,h)$ be the function given in Assumption 1, that is,

(A.1) $$U(C,h) = V_1(C) + \kappa V_2(h) + r\kappa V_1(C)V_2(h).$$

A necessary and sufficient condition for quasi-concavity of $U(C,h)$ is that

(A.2) $$D \equiv det\begin{pmatrix} 0 & U_1'(C,h) & U_2'(C,h) \\ U_1'(C,h) & U_{11}'(C,h) & U_{12}'(C,h) \\ U_2'(C,h) & U_{21}'(C,h) & U_{22}'(C,h) \end{pmatrix} \geq 0.$$

From (A.1) we obtain that

$$U_1'(C,h) = V_1'(C)(1 + r\kappa V_2(h)), \quad U_2'(C,h) = \kappa V_2'(h)(1 + rV_1(C)),$$

$$U_{11}'(C,h) = V_1''(C)(1 + r\kappa V_2(h)), \; U_{22}'(C,h) = \kappa V_2''(h)(1 + rV_1(C))$$

and

$$U_{12}'(C,h) = r\kappa V_1'(C)V_2'(h).$$

Hence, it follows that

(A.3) $$D = \kappa(1 + r\kappa V_2(h))(1 + rV_1(C))[2(V_1'(C)V_2'(h))^2 r\kappa - (V_1'(C))^2 V_2''(h)(1 + r\kappa V_2(h))$$
$$-\kappa(V_2'(h))^2 V_1''(C)(1 + rV_1(C))].$$

In order for utility to be strictly increasing in *C* and strictly decreasing in *h* is must be the case that

(A.4) $$1 + r\kappa V_2(h) > 0 \text{ and } 1 + rV_1(C) > 0.$$

Since $\kappa > 0$ it follows from (A.3) that when (A.4) holds then utility is quasi-concave provided

(A.5) $$2(V_1'(C)V_2'(h))^2 r\kappa - (V_1'(C))^2 V_2''(h)(1 + r\kappa V_2(h)) - \kappa(V_2'(h))^2 V_1''(C)(1 + rV_1(C)) \geq 0.$$

The inequality in (A.5) evidently holds when *r* is non-negative provided (A.4) holds. Moreover, (A.5) is equivalent to

(A.6) $$r\kappa \geq \frac{(V_1'(C))^2 V_2''(h) + \kappa(V_2'(h))^2 V_1''(h)}{2(V_1'(C)V_2'(h))^2 - V_2''(h)V_2(h)(V_1'(C))^2 - V_1''(C)V_1(C)(V_2'(h))^2}.$$

Since (A.5) holds for non-negative r when (A.4) holds the right hand side of (A.6) must be non-positive.

Q.E.D.

Statistics Norway

Postal address:
PO Box 8131 Dept
NO-0033 Oslo

Office address:
Akersveien 26, Oslo
Oterveien 23, Kongsvinger

E-mail: ssb@ssb.no
Internet: www.ssb.no
Telephone: + 47 62 88 50 00

ISSN: 1892-753X

Design: Siri Boquist

Statistisk sentralbyrå
Statistics Norway