Centre for Household, Income, Labour and Demographic economics

Working Papers

Evaluating Alternative Representations of the Choice Sets in Models of Labour Supply

R. Aaberge, U. Colombino and T. Wennemo

ChilD n. 14/2006

e-mail: de-child@unito.it
Web site: http://www.child-centre.it

Evaluating Alternative Representations of the Choice Sets in Models of Labour Supply

R. Aaberge, Statistics Norway, Oslo, Norway
U. Colombino, Department of Economics, Turin, Italy
T. Wennemo, Statistics Norway, Oslo, Norway

Abstract
During the last two decades, the discrete-choice modelling of labour supply decisions has become increasingly popular, starting with Aaberge et al. (1995) and van Soest (1995). Within the literature adopting this approach there are however two potentially important issues that so far have not been given the attention they might deserve. A first issue concerns the procedure by which the discrete alternatives are selected to enter the choice set. For example van Soest (1995) chooses (non probabilistically) a set of fixed points identical for every individual. This is by far the most widely adopted method. By contrast, Aaberge et al. (1995) adopt a sampling procedure suggested by McFadden (1978) and also assume that the choice set may differ across the households. A second issue concerns the availability of the alternatives. Most authors assume all the values of hours-of-work within some range [0, H] are equally available. At the other extreme, some authors assume only two or three alternatives (e.g. non-participation, part-time and full-time) are available for everyone. Aaberge et al. (1995) assume instead that not all the hour opportunities are equally available to everyone; they specify a probability density function of opportunities for each individual and the discrete choice set used in the estimation is built by sampling from that individual-specific density function. In this paper we explore by simulation the implications of

- the procedure used to build the choice set (fixed alternatives vs sampled alternatives)
- accounting or not accounting for a different availability of alternatives.

The results of the evaluation performed in this paper show that the way the choice set is represented have little impact on the fitting of observed values, but a more significant and important impact on the out-of-sample prediction performance.

Key words: Labour supply, discrete-choice models, quantity constraints, prediction performance

JEL: C51, C52, H31, J22

Acknowledgement: Parts of this paper was written when the author was visiting ICER in Torino. ICER is gratefully acknowledged for providing financial support and excellent working conditions.

1. Introduction

The idea of modelling labour supply decisions as discrete choices has become more and more popular during the last two decades. In this paper we examine through a simulation exercise an issue that has received much less attention than it might deserve: the implications of alternative methods of representing the choice set within the discrete choice approach.

The discrete choice approach has gained a prominent position as an outcome of the process aimed at solving or circumventing some theoretical and computational problems to be faced in micro-econometric research when analyzing choices subject to complicated opportunity constraints.

Let us consider the standard labour supply framework:

$$\begin{aligned} &\max U(h,x) \\ &s.t. \\ &x \leq wh + I \quad and \quad h \geq 0 \end{aligned} \tag{1.1}$$

where U is a deterministic utility function, x is consumption, h represents hours of work, w is the (constant) hourly wage rate and I is the exogenous income. Using the Kuhn-Tucker conditions associated to (1.1) – and assuming for simplicity an interior solution – under appropriate conditions one can obtain the optimal labour supply $h*$ as a function of w and I:

$$h* = h(w, I) \tag{1.2}$$

Then some empirical specification of $h(w, I)$ can be estimated and used for example to simulate the effects of policies implying changes in w and/or in I. The linear budget constraint in problem (1.1), however, very rarely corresponds to reality. Consider a well-known example: taxes and transfers on income in general imply a non-linear constraint. The budget constraint would then be:

$$x \leq wh + I - t(wh, I) \tag{1.3}$$

where t represents the tax-benefit rule that computes the taxes to be paid and the transfers to be received given gross incomes (wh, I). Taking (1.3) into account, we might still be able to characterize the optimal solution as a function of w and I ,

$$h* = h^t(w, I) \tag{1.4}$$

and estimate $h^t(w, I)$. However, $h^t(w, I)$ depends on the current tax-benefit rule t and therefore it cannot be used to simulate policies that introduce a different tax rule, say t'. The problem is that the behavioural function h^t in general mixes up preferences and constraints.[1] More generally, the opportunity set might be defined by complicated budget and quantity constraints that do not even allow recovering a closed form solution for $h*$. What we really need is an estimate of the utility function $U(h,x)$ itself. Once preferences are estimated, in principle we are able to simulate the effect of any policy by solving $\max U(h,x)$ subject to the appropriate constraints.

Heckman (1974) took full account of the non-linearity of the budget constraint in the estimation and simulation of microeconometric models. The problem addressed is the evaluation of a child related welfare policy that introduces significant complications in the budget set. Heckman proposed a particular method of recovering preferences by using the conditions to be fulfilled by the marginal rate of substitution for $h*$ to be located on a particular point of the budget set. Shortly after, a series of papers by Hausman and various co-authors proposed a method specifically addressed to piece-wise linear budget constraints (e.g. Hausman, 1979). Both Heckman (1974) and Hausman (1979) work through the implications of the Kuhn-Tucker conditions. The solution can be located in different ranges of values along the budget constraint. Corresponding to each possible range of values there is a condition involving the preference parameters. Choosing a convenient stochastic specification, we can express the probability that those various conditions alternatively hold, write down the sample likelihood and estimate the preference parameters. Useful presentations of this class of methods are provided by Moffit (1986) and Blundell and MaCurdy (2000).

Soon it emerged that the approach described above presents three main problems. First, it works well with convex budget sets (e.g. those generated by progressive taxation) and a two-good application (e.g. h and x in the individual labour supply application) but it tends to become computationally cumbersome when the agents face non-convex budget sets and when more than two goods are object to choice (e.g. when the agent is a many-person household). Second, in view of the computational problems, the above approach essentially forces the researcher to choose relatively simple specifications for the utility function or the labour supply functions Third – in the approach proposed by Hausman and associates – computational and statistical consistency of ML estimation of the model requires imposing a priori quasi-convexity of preferences (e.g. see MaCurdy et al., 1990).

Due to these emerging problems, applied researchers have started to make use of another innovative research effort also maturated in the first half of the 70's, i.e. the discrete choice modelling approach developed by McFadden (1974). As far as the labour supply application is concerned, the approach essentially consists in representing the budget set with a set of discrete 'points'. Let $[0,H]$ be the (continuous) range of possible values for hours of work h. Let us pick s points $h_1, h_2, ..., h_s$ to "represent" $[0,H]$. The utility level attained at point k is $U(x_k, h_k)$, where x_k is obtained through

[1] $h^t(w,I)$ is called "mongrel" labour supply function by Blomquist (1988).

some budget rule such as (1.4). Now let us assume that $U(x_k, h_k)$ is a random variable that can be decomposed additively into a systematic part containing the observable $v(x_k, h_k)$ and a random component e_k that accounts for the effect of unobservables: $U(x_k, h_k) = v(x_k, h_k) + e_k$. The assumption that the random term e_k is Type I Extreme Value i.i.d. leads to the well known multinomial logit expression for the probability that point j (i.e. the job with hours h_j) is chosen:[2]

$$P(j) = \Pr\left(U\left(x_j, h_j\right) = \max\left(U\left(x_1, h_1\right), ..., U(x_s, h_s)\right)\right) = \frac{\exp\left(v(x_j, h_j)\right)}{\sum_i \exp\left(v(x_i, h_i)\right)}. \tag{1.5}$$

The corresponding likelihood function can then easily be computed and maximised in order to estimate the parameters of the utility function. This approach is computationally very convenient when compared to the previous one, since it does not require going through complicated Kuhn-Tucker conditions involving derivatives of the utility function and of the budget constraints. As a consequence it is not affected by the complexity of the rule that defines the budget set or by how many goods are contained in the utility function. Equally important, the deterministic part of the utility function can be specified as very flexible without worrying for the computational problems.

During the last two decades, this approach has become increasingly popular, starting with Aaberge et al (1995) and van Soest (1995). Within the literature adopting this approach there are however two potentially important issues that are worthwhile analyzing in their implications and that so far have not been given the attention they might deserve.

A first issue concerns the procedure by which the discrete alternatives are included in the choice set. Most authors (e.g., among others, van Soest (1995), Duncan and Weeks (1997), Blundell, Duncan et al. (2000)), Kornstad and Thoresen (2004) choose (not probabilistically) a set of fixed points identical for every individual[3]. By contrast, Aaberge et al. (1995) and Aaberge et al (1999) adopt a sampling procedure originally proposed by McFadden (1978) and also assume that the choice set may differ across the households.

A second issue concerns the availability of the alternatives. Most authors assume all the values in $[0, H]$ - or in some discrete subset - are equally available. At the other extreme, some authors (e.g. Zabalza et al. (1980) assume only two or three alternatives (e.g. non-participation, part-time and full-time) are available for everyone. More generally, Aaberge et al. (1995, 1999, 2000, 2004) assume that the hour opportunities in $[0, H]$ are not equally available to everyone. More specifically, they assume that there is a probability density function of opportunities for each individual. The discrete choice set used in the estimation (and subsequently in the simulations) is built by sampling from that individual-specific density function.

[2] A random variable e has a (standard) Type I extreme value distribution if Prob($e \leq k$) = $\exp\left(-\exp(-k)\right)$.

[3] See Creedy and Kalb (2004) for a survey of discrete hours alternative approaches.

In what follows we explore by simulation the implications of

- the procedure used to build the choice set (fixed alternatives vs sampled alternatives)
- accounting vs not accounting for a different availability of alternatives (uniform availability vs heterogeneous availability)

upon the precision of the estimates and of policy simulation results.

The simulation is organized as follows. First, we use a previously estimated model as the "true" one. This model is characterized by heterogeneous availability of alternatives (across different hour values and among different individuals). Second, from the "population" described by the "true" model we generate a large sample of observations. Third, we use the data from this sample to estimate various versions of models that adopt the same specification of preferences as in the "true" model but differs in the way the choice set is represented (sampled vs fixed alternatives, number of alternatives, heterogeneous vs uniform availability of alternatives). Fourth, we investigate the effect of different specifications of the choice set upon the in-sample and out-of-sample prediction error. Last, we exemplify the trade-off between prediction precision and computational burden of different choice set specifications.

2. The "true" model

The "true" model is defined along the lines adopted in Aaberge et al. (1995) as well as in several successive papers.[4] The individuals maximise their utility by choosing among opportunities defined by hours of work, hourly wage and non-pecuniary attributes of the job. The utility is assumed to be of the following form

$$U\left(f(wh,I),h,j\right)=v(f(wh,I),h)+e(w,h,j) \quad (2.1)$$

where w is the wage rate, h is hours of work, I is exogenous income including the husband's labour income, f is a function that transforms gross income into income after tax, i.e. $f\left(wh,I\right)$ is disposable income (income after tax), j is a variable that captures other job characteristics and e is a random variable that is supposed to account for unobservables affecting tastes for a given job across individuals as well as across job opportunities for a given individual.[5] Commuting time or required skill are type of characteristics captured by j. The individual is supposed to choose a "job" from a choice set B that may differ across individuals. Each job alternative in B contains a wage rate w, hours of work h and unobserved (to the analyst) job characteristics such as environmental characteristics and skill content of the job. Moreover, B contains also non-market activities, i.e. jobs with $w=0$ and $h=0$.

[4] e.g. Aaberge, Colombino and Strøm, (1999, 2000, 2004) and Aaberge, Colombino, Strøm and Wennemo (2000).

By assuming that e is type I extreme value distributed and that the specification (2.1) is valid, it turns out that the probability density that opportunities with hours h and wage rate w are chosen has the following expression[6]

$$(2.2) \qquad j(h,w) \equiv \Pr\left[U(f(wh,I),h) = \max_{(x,y)\in B} U(f(xy,I),y)\right] = \frac{\exp(v(f(wh,I),h)p(h,w)}{\iint \exp(v(f(xy,I),y)p(x,y)dxdy}$$

where $p(h,w)$ is the density of choice opportunities which can be interpreted as the relative frequency (in the choice set) of opportunities with hours h and wage rate w. Opportunities with $h=0$ (and $w=0$) are non-market opportunities (i.e. alternative allocations of "leisure"). Thus, the density (2.2) will form the basis of estimating the parameters of the utility function and the choice sets.

In practice, the estimation adopts a discretized version of (2.2). Let $q(h,w)$ be some known joint density function (e.g. empirically fitted to the observations on h and w). Let us represent the latent choice set B with a sample S containing M points, where one is the chosen (observed) point and the other M-1 are sampled from $q(h,w)$. It can be shown (McFadden, 1978; Ben Akiva and Lerman, 1985) that consistent estimates of $v(f(wh,I),h)$ and $p(h,w)$ can still be obtained when (2.2) is replaced by

$$(2.3) \qquad \Pr\left[U(f(wh,I),h) = \max_{(x,y)\in S} U(f(xy,I),y)\right] = \frac{\exp\big(v(f(wh,I),h)\big)\, p(h,w)/q(h,w)}{\sum_{(x,y)\in S} \exp\big(v(f(xy,I),y)\big)\, p(x,y)/q(x,y)}.$$

We select a sample of married/cohabitating females. The systematic part of their utility function (2.1) is specified as follows

$$(2.4) \qquad \begin{aligned} v(h,w) &= a_2\left(\frac{f(wh,I)^{a_1}-1}{a_1}\right) + (a_4 + \\ &+ a_5 \log A + a_6(\log A)^2 + a_7 Ch_1 + a_8 Ch_2 + a_9 Ch_3)\left(\frac{L^{a_3}-1}{a_3}\right) \end{aligned}$$

where L is leisure, defined as $L = 1-(h/8736)$ and h is yearly hours of work, A is age and Ch_1, Ch_2 and Ch_3 are number of children below 3, between 3 and 6 and between 7 and 14 years old. In the specification of the probability density of opportunities $p(h,w)$ we will assume that offered hours and offered wages are independently distributed. The justification for this is that offered hours, in

---

[5] In most of the papers where the model is presented, the multiplicative specification is chosen, i.e. U = ve . We formulate here the model in the additive form in order to make it more easily comparable to similar models that appear in the literature.

particular normal working hours, are typically set in rather infrequent negotiations between employers and employees associations, while wage negotiations are far more frequent in which the hourly wage tend to be set independent of working hours. Thus, we specify the density of opportunities requiring h hours of work and paying hourly wage w as follows

$$(2.5) \qquad p(h,w) = \begin{cases} p_0 g_1(h) g_2(w) & \text{if} \quad h > 0 \\ 1 - p_0 & \text{if} \quad h = 0 \end{cases}$$

where p_0 is the proportion of market opportunities in the opportunity set, and g_1 and g_2 are respectively the densities of hours and wages, conditional upon the opportunity being a market job.

In view of the empirical specification it is convenient to divide both numerator and denominator by $1 - p_0$ and define $g_0 = \frac{p_0}{1 - p_0}$. We can then rewrite the choice density (2.2) as follows:

$$(2.6) \qquad j(h,w) = \frac{\exp\left(v(h,w)\right) g_0 g_1(h) g_2(w)}{\exp\left(v(0,0)\right) + \int\limits_{x>0} \int\limits_{y>0} \exp\left(v(x,y)\right) g_0 g_1(x) g_2(y) dxdy}$$

for $\{h, w\} > 0$, and

$$(2.7) \qquad j(0,0) = \frac{\exp\left(v(0,0)\right)}{\exp\left(v(0,0)\right) + \int\limits_{x>0} \int\limits_{y>0} \exp\left(v(x,y)\right) g_0 g_1(x) g_2(y) dxdy}$$

for $\{h, w\} = 0$.

Offered hours are assumed to be uniformly distributed except for possible peaks at half-time (18-20 weekly hours), and to full-time (37-40 weekly hours). Thus, g_1 is given by

[6] For the derivation of the choice density (2.2), see Aaberge et al. (1999). Note that (2.2) can be considered as a special case of the more general multinomial type of framework introduced by Ben-Akiva and Watanatada (1981) and Dagsvik (1994).

$$(2.8) \qquad g_1(h) = \begin{cases} g & \text{if} \quad h \in [1,17] \\ g\exp(p_1) & \text{if} \quad h \in [18,20] \\ g & \text{if} \quad h \in [21,36] \\ g\exp(p_2) & \text{if} \quad h \in [37,40] \\ g & \text{if} \quad h \in [41,H] \end{cases}$$

where H is the maximum observed value of h. Thus, this opportunity density for offered hours implies that it is far more likely to find jobs with hours that accord with full-time and standard part time positions than jobs with other working loads. Since the density values must add up to 1, we can also compute g according to:

$$(2.9) \qquad g = \left((17-1)+(20-18)\right)\exp(p_1)+(36-21)+(40-37)\exp(p_2)+(H-41)\right)^{-1}$$

Moreover we write

$$(2.10) \qquad g_0 = \exp(q_0).$$

In Table 2.1 we refer to p_1, p_2 *and* q_0 as the parameters of the *job opportunity density*.

The density of offered wages is assumed to be lognormal with mean that depends on length of schooling (*Ed*) and on past potential working experience (E*xp*), where experience is defined to be equal to age minus length of schooling minus five, i.e.

$$(2.11) \qquad \log w = b_0 + b_1 Exp + b_2 Exp^2 + b_3 Ed + s h$$

where h is standard normally distributed.

Using (2.8) and (2.10) we can write the choice density as follows:

$$(2.12) \qquad j(w,h) = \frac{\exp\left(v(f(wh,I),h)+(q_0+\ln g_2(w))d_0(h)+p_1 d_1(h)+p_2 d_2(h)\right)}{\int_x\int_y \exp\left(v(f(xy,I),y)+(q_0+\ln g_2(x))d_0(y)+p_1 d_1(y)+p_2 d_2(y)\right)dxdy}$$

where

$$(2.13) \qquad \begin{aligned} d_0(h) &= 1 \text{ if } h>0;\ 0 \text{ otherwise} \\ d_1(h) &= 1 \text{ if } h \in [18,20];\ 0 \text{ otherwise}\ . \\ d_2(h) &= 1 \text{ if } h \in [37,40];\ 0 \text{ otherwise} \end{aligned}$$

In what follows we will refer to $d_0(h)$ as the "job" dummy, since it captures the relative frequency of market opportunities to non-market opportunities; we will refer to $d_1(h)$ and $d_2(h)$ as the "peaks" dummies, since they are meant to capture the "peaks" in the density of hours corresponding to part-time and full-time jobs.

The estimation of the model is based on data for 1842 married/cohabitating females from the 1995 Norwegian Survey of Level of Living. We have restricted the ages of the females to be between 20 and 62 years in order to minimize the inclusion in the sample of individuals who in principle are eligible for retirement, since analysis of retirement decisions is beyond the scope of this study. Although the model adopted was originally developed for analysing simultaneous household partners' behaviour, we focus here on women's behaviour in order to simplify the execution and the interpretation of the simulation exercise. Husband's income as well as the couple's non-labour income are treated as exogenous and included in disposable income $f(wh, I)$.

The parameters appearing in expressions (2.3) – (2.10) are estimated by the method of maximum likelihood using the sampling procedure illustrated in expression (2.3). Each of the choice sets are represented by a set S that includes the observed choice plus 999 independent draws *(h, w)* from densities *q(w, h)* previously fitted to the observed values of w and h. If (w_s, h_s) are the observed values for a particular individual, the corresponding contribution to the likelihood function is:

$$(2.14)\; P(w_s, h_s \mid S) = \frac{v\left(f(w_s h_s, I), h_s\right)\exp\left(\left(q_0 + \ln g_2(w_s)\right) d_0(h_s) + p_1 d_1(h_s) + p_2 d_2(h_s) - \ln q(w_s, h_s)\right)}{\sum_{i \in S} v\left(f(w_i h_i, I), h_i\right)\exp\left(\left(q_0 + \ln g_2(w_i)\right) d_0(h_i) + p_1 d_1(h_i) + p_2 d_2(h_i) - \ln q(w_i, h_i)\right)}.$$

The estimates of the parameters of the opportunity density parameters and the parameters of the utility function are reported in Tables 2.1 and 2.2.

We then generate the sample that will be used in the simulation exercise. Given the estimates of Tables 2.1 and 2.2, for each one of the 1842 individuals in the sample we build a choice set containing 1000 alternatives. For each alternative i we compute the corresponding utility level by adding to the systematic part $v\left(f(w_i h_i, I), h_i\right)$ a value of e from the extreme value distribution of type I. The simulated chosen alternative is the one with the highest utility level. We make 6 draws of e so that for each individual we have 6 different choices, thus obtaining a total sample of 6×1842 = 11052 observations.

Table 2.1. Hours and wage densities, Norway 1994

	Parameter	Estimate	Std. Dev.
Job opportunity			
	q_0	-0.60	(0.10)
Hours			
Half-time	p_1	0.46	(0.10)
Full-time	p_2	1.57	(0.07)
Wage			
	b_0	0.24	(0.01)
	b_1	3.62	(0.05)
	b_2	2.41	(0.26)
	b_3	-3.67	(0.58)
	s	4.10	(0.35)

Table 2.2. Estimates of the parameters of the utility functions for married/cohabitating females. Norway 1994

Variable	Parameter	Estimate	Std. Dev.
Consumption			
	α_1	0.39	(0.11)
	α_2	4.42	(0.44)
Leisure			
	α_3	-4.57	(0.53)
	α_4	168.88	(27.47)
Log age	α_5	-94.29	(15.32)
Log age squared	α_6	13.35	(2.16)
Number of children below 3 years old	α_7	0.44	(0.23)
Number of children 3-6 years old	α_8	1.23	(0.24)
Number of children 7-14 years old	α_9	1.05	(0.19)

3. Alternative representations of the choice sets

3.1 Selection of alternatives

As we have already mentioned in the Introduction, the first issue in choice set representation concerns the procedure used to select the alternatives. In many applications, including labour supply modelling, the choice set contains a very large (or even infinite) number of alternatives. For instance, if we model couples labour supply and the decision period is the year, considering 1 hour intervals and 16 hours available during the day, there are $(16\times 365)^2 = 34{,}105{,}600$ alternatives. This would imply a very heavy computational burden, since for each alternative we must compute the couple's budget by applying a possibly complicated tax rule. Thus it is convenient to work with a smaller choice set somehow representative of the true one. Ben-Akiva and Lerman (1985) present a detailed treatment of the procedures that might be used when the number of alternatives contained in the choice set is very large (or even infinite) so that a complete enumeration is computationally too costly:

- Aggregation of alternatives
- Sampling of alternatives

The procedure consisting in selecting a fixed number of hours' values can be interpreted as an aggregation procedure. Instead of using all the possible values between 0 and T, the $(0,T)$ range is divided into sub-intervals and then the mid (or maybe the average) value of h in each interval is chosen to 'represent' all the values of that interval. The authors adopting this procedure realize that it introduces measurement errors, but tend to assume they are of minor importance. For example van Soest (1995) reports that some experiments with a different number of points did not show significant differences in parameter estimates, however a systematic investigation of the implication of that procedure has never been done either theoretically or empirically. However, if one interprets the approximation as an aggregation procedure, the analysis provided by Ben-Akiva and Lerman (1985) can be applied to clarify the issue.

We will assume the average of h in each sub-interval is chosen as representative (instead of the more common procedure of choosing the mid point: of course the two are very close and in fact coincide if the values of h are continuous or if each interval contains an uneven number of values). Let us define (we drop the subscript of the household to simplify the notation):

$$v_j \equiv v\left(f\,(wh_j\,, I), h_j \right). \tag{3.1}$$

Furthermore, let $\bar{v}^L \equiv \frac{1}{N^L}\sum_{j\in L} v_j$ = average systematic utility in sub-interval *L*, where N^L = number of elements in *L* and $\bar{h}^L$ = average value of h in sub-interval *L*. Ben-Akiva and Lerman show that the expected maximum utility attained on subinterval ℓ is

$$\hat{v}^\ell = \bar{v}^\ell + \ln(N^\ell) + \ln\left(D^\ell\right) \tag{3.2}$$

where $D^\ell \equiv \sum_j \exp\left(v_j^\ell - \bar{v}^\ell\right)\frac{1}{N^\ell}$. This last term is a measure of dispersion of v in sub-interval ℓ.

Accordingly, the probability that a value of h belonging to sub-interval L is chosen is

$$P(L) = \frac{\exp\left(\bar{v}^L + \ln(N^L) + \ln\left(D^L\right)\right)}{\sum_\ell \exp\left(\bar{v}^\ell + \ln(N^\ell) + \ln\left(D^\ell\right)\right)}. \tag{3.3}$$

To compare this with the expressions used in the fixed-alternatives approach it is useful to Taylor-expand v_j up to 2-order terms to get

$$P(L) \cong \frac{\exp\left(v\left(f(w\bar{h}^L, I), \bar{h}^L\right) + 0.5 s_{hh}^L v_{hh}^L + \ln(N^L) + \ln\left(D^L\right)\right)}{\sum_\ell \exp\left(v\left(f(w\bar{h}^\ell, I), \bar{h}^\ell\right) + 0.5 s_{hh}^\ell v_{hh}^\ell + \ln(N^\ell) + \ln\left(D^\ell\right)\right)} \tag{3.4}$$

where s_{hh}^ℓ is the variance of the values of h in sub-interval ℓ and v_{hh}^ℓ is the second (total) derivative of $v\left(f(w\bar{h}^\ell, I), \bar{h}^\ell\right)$ with respect to $\bar{h}^\ell$. It would be pointless to use (3.4) for estimation since it requires the very same computations that one wishes to avoid by aggregating alternatives. However (3.4) is useful in order to understand the type and the extent of the errors we incur by using various approximations. The expression typically used in the literature is:

$$\hat{P}(L) = \frac{\exp\left(v\left(f(w\bar{h}^L, I), \bar{h}^L\right)\right)}{\sum_\ell \exp\left(v\left(f(w\bar{h}^\ell, I), \bar{h}^\ell\right)\right)}. \tag{3.5}$$

Clearly, in expression (3.5) all the terms $0.5 s_{hh}^\ell v_{hh}^\ell + \ln(N^\ell) + \ln\left(D^\ell\right)$ are dropped. If these terms were equal across all the sub-intervals they would cancel out from (3.4) and (3.5) would be exact. In general however they will not be equal, and dropping them will lead to biased estimates. Nonetheless there are ways by which we could improve upon (3.5) when adopting aggregation as an approximation strategy, which however has never been considered in the literature on labour supply modelling:

- The dimension of N^ℓ of the sub-intervals - when not equal for all of them - is typically known and can be explicitly accounted for;
- s_{hh}^ℓ can also be computed;

- Depending on the functional form used for the utility function, the term v_{hh}^{ℓ} might be explicitly evaluated and accounted for;
- The terms $\ln\left(D^{\ell}\right)$ in general will vary both across sub-intervals and across individuals; however we might capture at least some of their effect by introducing a set of dummies (as many as the number of sub-intervals - 1).

Summing up, the aggregation of alternatives implies biased estimates. The bias could be moderated by using various possible corrections suggested by expression (3.4) itself. Up to now, however, it must be said that the literature on labour supply has treated this issue in a rather superficial way (when compared, for instance to the literature on transportation or location choices).

Sampling of alternatives, on the other hand, offers the possibility of working with a relatively small choice set and at the same time preserving the consistency of the estimates. The basic results were established by McFadden (1978). Ben-Akiva and Lerman (1985) also provide a very useful and more practically oriented survey, together with some additional theoretical results. Let us represent the true choice set B with a sample S containing M points, where one is the chosen (observed) point and the other M-1 are sampled from $q(h,w)$. Let q_i be the probability of sampling point h_i. It can be shown (McFadden, 1978; Ben Akiva and Lerman, 1985) that consistent estimates of $v(f(wh,I),h)$ and $p(h,w)$ can still be obtained when the true choice set B is replaced by S and the probability of observing choice j is evaluated as follows:

$$(3.6) \qquad P(j|S) = \frac{\exp\left(v(f(wh_j,I),h_j) - \ln(q_j)\right)}{\sum_{h_i \in S} \exp\left(v(f(wh_i,I),h_i) - \ln(q_i)\right)}.$$

If a simple random sampling is adopted, all the *q's* are equal and cancel out. Typically more sophisticated sampling procedures are used since they are expected to be more efficient. For instance, a common procedure consists in using as sampling probabilities the observed relative frequencies of choice possibly differentiated according to personal characteristics of the decision units. Besides Ben-Akiva and Lerman (1985), also Train et al. (1987) present a very detailed application of this procedure.

3.2 Availability of alternatives

A second and possibly even more substantial issue is whether or not account is taken of the different availability of job-types on the market. Some authors have made the extreme choice of assuming the choice set contains only two or three alternatives (e.g. non-participation, part-time and full-time). More common, however, is the approach of choosing a few equally spaced points in the interval [0,H], without taking into account the possibility that some type of opportunities maybe more easily available than others. Other authors (Aaberge et al. 1995, 1999, 2004) do account for this possibility as well as for the relative density of jobs as a function of personal characteristics (see Section 2). In practice, their specification boils down to "augmenting" the term v with a set of appropriately defined dummy variables. Also van Soest (1995) introduces similar dummies, although he gives them a different interpretation in terms of utility costs or *premia* attached to some range of hour values.[7]

3.3 The simulation exercise

In what follows we use the sample generated according to the true model to estimate various versions of models generated according to the various possible representations of the choice set as discussed above. In these models the wage rate is kept fixed for each individual, i.e. it does not vary across alternatives as it is allowed for in the "true" model; moreover it is simultaneously estimated as in (2.12), instead we use a wage equation to predict the wage for non-participants.[8] This simplification is introduced in order to make the simulation results more easily interpretable.

The more general versions of the models are

$$P(w_s,h_s|S)=\frac{\exp\big(v\big(f(w_sh_s,I),h_s\big)+q_0d_0(h_s)+p_1d_1(h_s)+p_2d_2(h_s)-\ln q(w_s,h_s)\big)}{\sum_{i\in S}\exp\big(v\big(f(w_ih_i,I),h_i\big)+q_0d_0(h_i)+p_1d_1(h_i)+p_2d_2(h_i)-\ln q(w_i,h_i)\big)} \tag{3.7}$$

when sampled alternatives are used, and

$$P(w_s,h_s|R)=\frac{\exp\big(v\big(f(w\,h_s,I),h_s\big)+q_0d_0(h_s)+p_1d_1(h_s)+p_2d_2(h_s)\big)}{\sum_{i\in R}\exp\big(v\big(f(w\,h_i,I),h_i\big)+q_0d_0(h_i)+p_1d_1(h_i)+p_2d_2(h_i)\big)} \tag{3.8}$$

when fixed alternatives are used. The dummies $d_0(h_i)$ and $(d_1(h_i),d_2(h_i))$ are defined as in (2.13). Dropping the job dummy $d_0(h_i)$ and/or the peaks dummies $(d_1(h_i),d_2(h_i))$ generates a more restrictive version of the model. The choice sets S and R contain alternatively 6 or 24 points. Altogether with have 16 models resulting from the combinations of the following possibilities:

1. *alternative generation*: fixed or sampled;

---

[7] There is still another approach, the so called Dogit model (Gaudry and Dagenais, 1979), to represent a non-uniform availability of alternatives. It is a generalization of the logit model, where the decision-maker may – with a given probability – be "captive" to one of the alternatives or otherwise choose freely from the whole choice set. The Dogit model has been recently used by Harris and Duncan (2002) in a labour supply application. We do not consider the Dogit model in the simulation exercise presented here.

[8] The estimates of the wage equation are available upon request from the authors.

2. *number of alternatives*: 6 or 24;
3. *job dummy*: included or dropped;
4. *peaks dummies*: included or dropped.

In the Tables that follow the 16 models are labelled as in Table 3.1. The parameter estimates of the 16 models are reported in the Appendix.

We are interested in the prediction performance of the models, both in-sample and out-of-sample (prediction of policy effects). Clearly, we expect the more flexible models (such as Id or IId) to perform better than more restrictive models. Also, we know that the models based on sampled alternatives are expected to produce consistent estimates, while those based on fixed alternatives are not. Therefore what in fact we want to explore is *how much better* the more flexible models do and *how much better* the models based on sampled alternatives perform. Finally, we also attempt to illustrate the trade-off between prediction performance and computational burden.

Table 3.1 Types of models

	Fixed alternatives				Sampled alternatives			
Model	Model Ia	Model Ib	Model Ic	Model Id	Model IIa	Model IIb	Model IIc	Model IId
Job dummy	*No*	*Yes*	*No*	*Yes*	*No*	*Yes*	*No*	*Yes*
Peaks dummies	*No*	*No*	*Yes*	*Yes*	*No*	*No*	*Yes*	*Yes*
Number of alternatives	*6*	*6*	*6*	*6*	*24*	*24*	*24*	*24*

4. Evaluation of the different modelling approaches

In order to evaluate the impact of alternative representations of the choice set on the performance of the models we proceed in the following way. First, for each of the 16 models we predict participation rates, hours of work and disposable income. The predictions are obtained individual by individual, evaluating the utility function – including the stochastic component drawn from the Type I extreme value distribution – at each alternative and identifying the selected alternative as the one with the highest utility level. The individual predictions are then aggregated into the 10 means of the 10 income deciles. We introduce the following summary measure of prediction performance z_k for model k,

(4.1) $$z_k = \sqrt{\sum_{j=1}^{10}\left(\frac{(\tilde{y}_{kj} - y_j)}{y_j}\right)^2}, \text{ k=1, 2 …,16,}$$

where y_j and $\tilde{y}_{kj}$ denote the outcomes in decile j of the true model and alternative model k, respectively. The outcomes are alternatively defined to be the job participation rate, hours of work and disposable income after tax. Next, we carry out a regression analysis where z is treated as a response variable and the following variables are treated as co-variates,

x_1 = 1 if the choice alternatives are sampled (= 0 if the choice alternatives are fixed),

x_2 = 1 if the number of choice alternatives is equal to 24 (= 0 if the number of alternatives is equal to 6),

x_3 = 1 when it is accounted for job entry (= 0 when it is *not* accounted for job entry),

x_4 = 1 when it is accounted for part-time and full-time peaks (= 0 when it is *not* accounted for part-time and full-time peaks).

The following equation forms the basis of the evaluation of alternative modelling approaches,

(4.2) $$z = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4 + a_{34}(x_3 * x_4)$$

where the coefficients a will measure the relevance of the different ways of specifying the choice set. Since the most important application of labour supply models is the evaluation of tax and welfare policy reforms, we focus on the prediction performance under alternative tax regimes. Namely, the steps above are repeated twice:

- Prediction of the outcomes under the current tax regime
- Prediction of the outcomes after the introduction of a flat tax (keeping total tax revenue constant).

4.1. Outcomes under the current tax regime

Tables 4.1 – 4.3 illustrate the results of the exercise under the current tax regime. Tables 4.1 and 4.2 refer to the eight models with fixed alternatives. In order to simplify the illustration we limit ourselves to the models without job and peaks dummies and to the models with both types of dummies. For each of the models and each of the 10 income deciles, we report the predictions of participation rates and hours of work in Table 4.1 and of after tax disposable income in Table 4.2. We do not report here the analogous results for the models with sampled alternatives, since they are very close to those with fixed alternatives. Even a causal inspection of the tables suggests that the prediction performance is pretty good whatever the model considered. Possibly the only entries where there seems to be some substantial error depending on the model used are the predictions of outcomes for the first decile. In any case, in order to systematically assess the impact of the characteristics of all the 16 models we run the regression (4.2) and report the results in Table 4.3.

Table 4.1 Prediction of participation rates and hours of work under the 1994 tax system. Fixed-alternatives models

Deciles	True model		Model Ia		Model Id		Model IIa		Model IId	
	Participation rates (per cent)	Annual hours of work	Participation rates (per cent)	Annual hours of work	Participation rates (per cent)	Annual hours of work	Participation rates (per cent)	Annual hours of work	Participation rates (per cent)	Annual hours of work
1	**58**	**568**	55	627	*43*	*514*	87	733	*55*	*568*
2	**65**	**715**	73	818	*61*	*730*	93	837	*67*	*708*
3	**79**	**937**	81	1000	*71*	*890*	95	989	*79*	*941*
4	**86**	**1157**	87	1179	*80*	*1130*	97	1125	*85*	*1153*
5	**91**	**1389**	92	1375	*87*	*1397*	96	1276	*90*	*1352*
6	**93**	**1527**	94	1494	*91*	*1541*	98	1429	*93*	*1528*
7	**93**	**1606**	95	1638	*91*	*1650*	99	1598	*94*	*1631*
8	**94**	**1695**	94	1701	*92*	*1735*	98	1667	*93*	*1672*
9	**94**	**1757**	95	1812	*93*	*1838*	99	1746	*96*	*1771*
10	**88**	**1523**	89	1631	*83*	*1566*	97	1676	*87*	*1567*
Mean	**84**	**1287**	86	1327	*79*	*1299*	96	1308	*84*	*1289*

Table 4.2 Prediction of disposable income (in NOK) under the 1994 tax system. Fixed-alternatives models

Deciles	True model	Model Ia	Model Id	Model IIa	Model IId
1	**168915**	170648	*169098*	171945	*168690*
2	**216080**	217801	*215357*	219415	*216333*
3	**244914**	245504	*243740*	245176	*243672*
4	**268880**	268308	*267340*	267880	*267659*
5	**290441**	290083	*290556*	288798	*289893*
6	**312088**	312113	*313719*	310410	*312446*
7	**336247**	335829	*337305*	334374	*336148*
8	**363833**	364607	*365453*	362513	*363739*
9	**403513**	405063	*405654*	403401	*404046*
10	**600841**	605283	*602163*	608705	*604516*
Mean	**320575**	321524	*321038*	321262	*320714*

Table 4.3. Estimates of equation (4.2) outcomes under the current tax regime

Outcome variable	α_0	α_1	α_2	α_3	α_4	α_{34}	R^2
Probability of participation	***.406 (.123)***	-.075 (.100)	.126 (.100)	-.266 (.142)	.086 (.142)	-.076 (.201)	.54
Hours of work	***.266 (.057)***	-.023 (.046)	.070 (.046)	***-.178 (.007)***	-.004 (.007)	.031 (.092)	.60
Income after tax	***.021 (.008)***	-.002 (.004)	***.007 (.004)***	***-.009 (.006)***	-.002 (.006)	-.003 (.008)	.50

*Standard deviation in parentheses

The results of Table 4.3 confirm the message conveyed by Tables 4.1 and 4.2. Very few coefficients are significant at standard levels (the significant ones are in bold italics). Overall one can conclude that there is little evidence of an important impact of alternative modes of representing the choice set as long as the replication of current values is concerned.

4.2 Outcomes under a Flat Tax reform.

In this second part of the simulation exercise, the models are run as after a hypothetical tax reform. Namely, a fixed proportional tax (Flat Tax) replaces the current tax system. The flat tax is determined running recursively the "true" model until the total tax revenue is the same as under the current system. Next, the "true" outcomes (hours and net disposable income) are compared to the outcomes simulated by the 16 models and the corresponding values of the z_k are computed. Tables 4.4 and 4.5 are analogous to Tables 4.1 and 4.2. Tables 4.6 and 4.7 replicate Tables 4.4 and 4.5, but with sampled-alternatives models. When it comes to reform simulations rather than current values replication, the differences in outcomes are somewhat more marked, and this is confirmed by Table 4.8 where eq. (4.2) is estimated, analogous to Table 4.3, but with reference here to post-Flat-Tax outcomes. There is a rather clear pattern of the effects of different modelling strategies in particular on the prediction of disposable income. For example, using 24 alternatives instead of 6 reduces the average percentage error by 0.8%. Using sampled-alternatives instead of fixed alternatives reduces it by 1.9%, introducing job and peaks dummies reduces it by 3.2%. Moreover, the detailed information provided by Tables 4.4 - 4.7 demonstrates that the less satisfactory out-of-sample prediction performance arises from discrepancies between the lower parts of the predicted and "observed" flat-tax distributions of hours of work and disposable income.

Table 4.4 Prediction of participation rates and hours of work under a flat tax reform. Fixed-alternatives models

Deciles	True model		Model Ia		Model Id		Model IIa		Model IId	
	Participation rates (per cent)	Annual hours of work	Participation rates (per cent)	Annual hours of work	Participation rates (per cent)	Annual hours of work	Participation rates (per cent)	Annual hours of work	Participation rates (per cent)	Annual hours of work
1	**69**	**987**	62	835	*55*	*826*	89	946	*63*	*890*
2	**75**	**1022**	77	943	*68*	*966*	95	1041	*74*	*943*
3	**84**	**1160**	83	1100	*76*	*1117*	96	1145	*83*	*1134*
4	**89**	**1315**	89	1260	*83*	*1279*	97	1271	*87*	*1291*
5	**93**	**1491**	93	1432	*89*	*1488*	97	1392	*91*	*1459*
6	**94**	**1609**	94	1542	*92*	*1626*	98	1543	*93*	*1609*
7	**94**	**1659**	95	1677	*92*	*1717*	99	1685	*94*	*1670*
8	**95**	**1742**	94	1735	*92*	*1786*	98	1727	*93*	*1720*
9	**95**	**1794**	96	1843	*94*	*1898*	99	1811	*96*	*1821*
10	**88**	**1549**	89	1647	*84*	*1619*	97	1721	*88*	*1606*
Mean	**88**	**1487**	87	1401	*82*	*1432*	96	1428	*86*	*1414*

Table 4.5 Prediction of disposable income (in NOK) under a flat tax reform. Fixed-alternatives models

Deciles	True model	Model Ia	Model Id	Model IIa	Model IId
1	**194076**	171081	*177612*	173092	*177934*
2	**234263**	214268	*220564*	222704	*220524*
3	**259189**	242704	*250457*	247374	*248492*
4	**279624**	266384	*272361*	271441	*271579*
5	**301124**	289038	*294062*	293453	*294681*
6	**323777**	314124	*320755*	319278	*319492*
7	**350809**	342509	*349310*	346358	*344397*
8	**383958**	375740	*379893*	378941	*377972*
9	**431297**	426513	*431747*	430622	*428668*
10	**651815**	649764	*651885*	657771	*652667*
Mean	**340993**	329213	*334865*	334103	*333641*

Table 4.6 Prediction of participation rates and hours of work under a flat tax reform. Sampled-alternatives models

Deciles	True model		Model Ia		Model Id		Model IIa		Model IId	
	Participation rates (per cent)	Annual hours of work	Participation rates (per cent)	Annual hours of work	Participation rates (per cent)	Annual hours of work	Participation rates (per cent)	Annual hours of work	Participation rates (per cent)	Annual hours of work
1	**69**	**987**	76	915	*65*	*883*	76	921	*65*	880
2	**75**	**1022**	83	982	*74*	*993*	84	985	*75*	992
3	**84**	**1160**	90	1159	*83*	*1131*	90	1151	*83*	1133
4	**89**	**1315**	92	1288	*88*	*1330*	93	1307	*89*	1338
5	**93**	**1491**	94	1449	*91*	*1493*	94	1460	*91*	1485
6	**94**	**1609**	95	1580	*94*	*1650*	95	1579	*94*	1646
7	**94**	**1659**	95	1671	*93*	*1691*	96	1675	*93*	1695
8	**95**	**1742**	97	1759	*96*	*1775*	97	1771	*96*	1774
9	**95**	**1794**	98	1806	*96*	*1811*	98	1807	*96*	1814
10	**88**	**1549**	92	1606	*88*	*1587*	92	1617	*88*	1586
Mean	**88**	**1487**	91	1422	*87*	*1434*	91	1427	*87*	*1434*

Table 4.7 Prediction of disposable income (in NOK) under a flat tax reform. Sampled-alternatives models

Deciles	True model	Model Ia	Model Id	Model IIa	Model IId
1	**194076**	175360	*178959*	175829	*178558*
2	**234263**	221008	*223384*	220745	*222943*
3	**259189**	248332	*249373*	247584	*249304*
4	**279624**	272276	*275414*	273516	*275739*
5	**301124**	293241	*296123*	293368	*295567*
6	**323777**	318317	*321883*	318698	*321400*
7	**350809**	346147	*348328*	346124	*348868*
8	**383958**	377469	*379296*	378295	*378984*
9	**431297**	430380	*430587*	429954	*431015*
10	**651815**	651514	*650805*	652383	*650766*
Mean	**340993**	333404	*335415*	333650	*335314*

Table 4.8 Contributions to the prediction performance: outcomes under a flat tax reform

Outcome variable	a_0	a_1	a_2	a_3	a_4	a_{34}	R^2
Probability of participation	***.215*** ***(.008)***	-.090 (.064)	.060 (.062)	-.110 (.091)	.077 (.091)	-.073 (.129)	.47
Hours of work	***.093*** ***(.013)***	***-.061*** ***(.010)***	***-.023*** ***(.010)***	-.003 (.014)	-.001 (.014)	.012 (.020)	.82
Income after tax	***.128*** ***(.006)***	***-.019*** ***(.005)***	***-.008*** ***(.005)***	***-.029*** ***(.007)***	***-.039*** ***(.007)***	***.036*** ***(0.01)***	.84

*Standard deviation in parentheses

4.4 The trade-off between prediction performance and computational costs

The different representations of the choice set imply different computational burdens, particularly with regards to the number of alternatives and to the procedure used to generate the alternatives. Depending on the availability of computing resources and time, the advantages of the various approaches to represent the choice set should be balanced against the computational costs. Table 4.9 reports the relative elapsed time (= 1 for the simplest model[9]) of a typical estimation run with four different type of models: fixed vs sampled alternatives and 6 vs 24 alternatives (accounting or not accounting for job and peaks dummies does not make any significant difference in terms of computation time). Table 4.10 summarizes the trade-off between relative computational burden and relative prediction error when predicting the effects of the tax reform. The relative prediction error of model k is z_k / z_*, where z_k is defined as in (4.1) and z_* is the lowest value among $z_1, z_2, ..., z_{16}$. Therefore, the relative prediction error is equal to 1 for the model with the lowest z_k.[10] Table 4.10 shows, for example, that going from the simplest and more restrictive choice set specification to the more complex one entails increasing the computational burden by a factor of 8.5, whilst the benefits are a reduction by factor of 10.3 of the relative prediction error of hours of work and a reduction by a factor of 2.18 of the relative prediction error of net available income.

Table 4.9 Relative computation time (estimation) for different models

	6 alternatives	24 alternatives
Fixed alternatives	1	4.62
Sampled alternatives	6.70	8.46

[9] The absolute computing time for estimating the simplest model was 2.42 seconds on a Alpha ES45, 1 Gb Mhz, 8 Gb workspace memory

[10] The values of z_* are 0.009 for the prediction of hours of work and 0.059 for the prediction of net disposable income.

Table 4.10 Summary of prediction performance (policy effect prediction) and computational burden in estimation of different types of models

Choice set representation	*Fixed Alternatives*								*Sampled Alternatives*							
Job dummy	*No*	*Yes*	*No*	*Yes*	*No*	*Yes*	*No*	*Yes*	*No*	*Yes*	*No*	*Yes*	*No*	*Yes*	*No*	*Yes*
Peaks dummies	*No*	*No*	*Yes*	*Yes*	*No*	*No*	*Yes*	*Yes*	*No*	*No*	*Yes*	*Yes*	*No*	*No*	*Yes*	*Yes*
Number of alternatives	*6*	*6*	*6*	*6*	*24*	*24*	*24*	*24*	*6*	*6*	*6*	*6*	*24*	*24*	*24*	*24*
Relative prediction error: hours of work	***10.3***	***10.3***	***10.3***	***10.3***	***7.8***	***7.8***	***7.8***	***7.8***	***3.6***	***3.6***	***3.6***	***3.6***	***1.0***	***1.0***	***1.0***	***1.0***
Relative prediction error: income after tax	***2.2***	***2.0***	***1.5***	***1.6***	***2.0***	***1.9***	***1.47***	***1.5***	***1.9***	***1.7***	***1.9***	***1.1***	***1.7***	***1.6***	***1.0***	***1.0***
Relative computational burden	***1.0***	***1.0***	***1.0***	***1.0***	***4.6***	***4.6***	***4.6***	***4.6***	***6.7***	***6.7***	***6.7***	***6.7***	***8.5***	***8.5***	***8.5***	***8.5***

5. Conclusions

We have performed a series of simulation exercises aimed at exploring the performance of different versions of a labour supply model, where different approaches to represent choice sets are used. The various models are estimated using a large sample generated by a "true" model, to which they can then be compared. In evaluating the models, we focus upon their ability, replicate the "true" outcomes under different tax regimes. It turns out that as far as the replication of the current-tax-regime outcomes are concerned, there is little evidence for important effects of alternative choice-set-representation procedures. Not even the number of alternatives contained in the choice set seems to matter. All the models predict very well, although there are some indications favouring the sampled-alternatives procedure. However, when it comes to predicting the effect of a flat-tax reform, the indications are more clear-cut: using sampled alternatives and accounting for heterogeneity of opportunities seems to significantly reduce the prediction errors. Clearly the sampled-alternative procedure is more costly computationally, so the benefits should eventually be balanced against the increased computational costs.

The prediction performance of current values does not discriminate between different models but the prediction performance of post-reform does: these results convey the important message that the ability of a model to replicate observed outcomes is not very informative. Ultimately, the models should be judge in their ability to do the job they are mainly built for, i.e. predicting the outcomes of policy changes.

Appendix

Here we report the parameter estimates of the true model and of the 16 alternative models.

Table A.1. Fixed-alternatives models Ix and IIx (x = a, b, c)

Variable	Para-meter	True model	Model Ia	Model IIa	Model Ib	Model IIb	Model Ic	Model IIc	Model Id	Model IId
Consumption										
	α_1	**0.39**	0.35	0.54	0.43	0.46	0.43	0.50	0.43	0.44
	α_2	**4.42**	2.46	3.70	3.97	4.55	4.05	4.64	4.17	4.38
Leisure										
	α_3	**-4.57**	-7.53	-3.18	-7.31	-6.72	-2.07	-0.14	-3.99	-4.15
	α_4	**168.88**	54.20	184.85	64.76	92.39	232.99	351.30	156.91	171.12
Log age	α_5	**-94.29**	-30.46	-102.83	-36.27	-51.64	-128.78	-193.30	-87.38	-95.45
Log age squared	α_6	**13.35**	4.32	14.62	5.15	7.33	18.27	27.48	12.40	13.54
Number of children below 3 years old	α_7	**0.44**	0.13	0.51	0.13	0.19	0.61	0.95	0.38	0.40
Number of children 3-6 years old	α_8	**1.23**	0.48	1.68	0.53	0.76	1.86	2.99	1.25	1.40
Number of children 7-14 years old	α_9	**1.05**	0.40	1.37	0.44	0.62	1.53	2.47	1.04	1.14
Job dummy	q_0	**-0.60**	-	-	-1.08	-2.33	-	-	-0.78	-2.10
Part-time dummy	p_1	**0.46**	-	-	-	-	-0.23	0.14	0.15	0.28
Full-time dummy	p_2	**1.57**	-	-	-	-	0.99	1.53	0.78	1.19

Table A.1. Sampled-alternatives models Ix and IIx (x = a, b, c)

	Para-meter	True model	Model Ia	Model IIa	Model Ib	Model IIb	Model Ic	Model IIc	Model Id	Model IId
Consumption										
	α_1	**0.39**	0.54	0.55	0.53	0.54	0.55	0.55	0.52	0.53
	α_2	**4.42**	3.96	3.93	4.72	4.64	4.56	4.51	4.70	4.62
Leisure										
	α_3	**-4.57**	-5.15	-5.27	-5.94	-6.10	-2.40	-2.49	-3.52	-3.60
	α_4	**168.88**	125.90	121.50	112.19	106.31	234.88	231.26	195.26	190.72
Log age	α_5	**-94.29**	-70.17	-67.75	-62.54	-59.28	-129.94	-128.03	-108.43	-105.95
Log age squared	α_6	**13.35**	9.96	9.62	8.88	8.42	18.46	18.19	15.39	15.04
Number of children below 3 years old	α_7	**0.44**	0.33	0.30	0.25	0.23	0.66	0.58	0.50	0.44
Number of children 3-6 years old	α_8	**1.23**	1.07	1.05	0.91	0.87	1.94	1.95	1.56	1.57
Number of children 7-14 years old	α_9	**1.05**	0.88	0.88	0.75	0.73	1.61	1.65	1.29	1.33
Job dummy	q_0	**-0.60**	-	-	-0.88	-0.86	-	-	-0.63	-0.60
Part-time dummy	p_1	**0.46**	-	-	-	-	0.44	0.44	0.53	0.52
Full-time dummy	p_2	**1.57**	-	-	-	-	1.66	1.63	1.56	1.54

References

Aaberge, R.., J.K. Dagsvik and S. Strøm (1995): "Labor Supply Responses and Welfare Effects of Tax Reforms", *Scandinavian Journal of Economics*, **97**, 4, 635-659.

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

Aaberge, R., U. Colombino and S. Strøm (2000): "Labour Supply Responses and Welfare Effects from Replacing Current Tax Rules by a Flat Tax: Empirical Evidence from Italy, Norway and Sweden", *Journal of Population Economics,* **13**, 595-621.

Aaberge, R., Colombino U., Strøm S. and T. Wennemo (2000): "Joint Labour Supply of Married Couples: Efficiency and Distribuitional Effects of Tax Reforms", in: Sutherland, Mitton, Sutherland and Weeks (eds.) *Microsimulation Modelling for Policy Analysis: Challenges and Innovations*, Cambridge University, Cambridge, U.K.

Aaberge, R., U. Colombino and S. Strøm (2004): "Do More Equal Slices Shrink the Cake? An Empirical Investigation of Tax-Transfer Reform Proposals in Italy", *Journal of Population Economics*, **17**, 4.

Ben-Akiva, M. and T. Watanatada (1981): "Application of a Continuous Spacial Choice Logit Model", in Manski, C. F. and McFadden D. (eds.) *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press.

Ben-Akiva, M., and S. R. Lerman (1985): *Discrete choice analysis*, (MIT Press, Cambridge).

Blomquist, S. (1988): "Non-linear Taxes and Labor Supply", *European Economic Review*, **32**, 1213-1226.

Blundell, R. and T. MaCurdy (2000): "Labor Supply: A Review of Alternative Approaches", in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics* , Elsevier North-Holland.

Blundell, R., A. Duncan, A., J. McCrae and C. Meghir (2000): "The Labour Market Impact of the Working Families' Tax Credit", *Fiscal Studies*, **21**, 75-100

Creedy, J. and G. Kalb (2005): "Discrete Hours Labour Supply Modelling: Specification, Estimation and Simuation", *Journal of Economic Surveys*, **19**, 97-734

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

Duncan, A. and M. Weeks (1997): "Behavioural Tax Microsimulation with Finite Hours Choices", *European Economic Review*, **41**, 619-626.

Gaudry, M. and M. Dagenais (1979): "The Dogit Model", *Transportation Research*, **13B**, 105 – 112.

Harris, M.N. and A. Duncan (2002): "Intransigencies in the Labour Supply Choice", Melbourne Institute Working Paper No 17/02.

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

Heckman, J. (1974): "Effects of Child-Care Programs on Women's Work Effort", *Journal of Political Economy*, **82**, 136-163.

Kornstad, T. and T.O. Thoresen (2004): "Means-testing the Child Benefit", *Review of Income and Wealth*, **50**, 29-49.

MaCurdy, T., D. Green and H. Paarsch (1990): "Assessing Empirical Approaches for Analyzing Taxes and Labor Supply", *Journal of Human Resources,* **25**, 415–49.

McFadden, D. (1974): "Conditional Logit Analysis of Qualitative Choice Behavior", in P. Zarembka (ed.), *Frontiers in Econometrics,* Academic Press, New York.

McFadden, D. (1978): "Modelling the Choice of Residential Location" in A. Karlquist, L. Lundquist, F. Snickard and J.J. Weilbull (eds.): *Spatial Interaction Theory and Planning Models,* Amsterdam, North-Holland.

Moffitt, R. (1986): "The Econometrics of Piecewise-Linear Budget Constraints: A Survey and Exposition of the Maximum Likelihood Method", *Journal of Business & Economic Statistics*, **4**, 317-28.

Train, K.E., D. L. McFadden and M. Ben-Akiva (1987): "The demand for local telephone service", *Rand Journal of Economics*, **18**, 109-123.

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

Zabalza, A., C. Pissarides and M. Barton (1980): "Social security and the choice between full-time work, part-time work and retirement", *Journal of Public Economics*, **14**, 245-276.