Discussion Papers No. 265, January 2000
Statistics Norway, Research Department

Yun Li

Modeling the Choice of Working when the Set of Job Opportunities is Latent

Abstract:
In this paper we analyze the decision of "working" versus "not working" within a discrete choice framework, where number of available jobs is confined and related to individual characteristics. In this way the market constraint from the demand side is taken into account. We also accommodate the notion of job specific non-pecuniary attributes in the specification of preferences. We apply panel data to estimate the model. To this end a particular estimation method is developed that accounts for possible serial dependence in the preferences.

Keywords: Labor supply, Discrete choice, Logit model

JEL classification: C25, J22

Acknowledgement: John K, Dagsvik has proposed the particular modeling framework applied in this paper. I am grateful for his guidance, comments and support.

Address: Yun Li, Statistics Norway, Research Department.

Discussion Papers comprise research papers intended for international journals or books. As a preprint a Discussion Paper can be longer and more elaborate than a standard journal article by including intermediate calculation and background material etc.

Abstracts with downloadable PDF files of
Discussion Papers are available on the Internet: http://www.ssb.no

For printed Discussion Papers contact:

Statistics Norway
Sales- and subscription service
N-2225 Kongsvinger

Telephone: +47 62 88 55 00
Telefax: +47 62 88 55 95
E-mail: Salg-abonnement@ssb.no

1. Introduction

There is a huge literature on labor supply of married women (see Killingsworth and Heckman (1986) for a survey). In most studies, the structure of modeling framework rests upon the assumption that the fundamental choice variables in this context are "consumption" (composite) and "leisure" (hours of work), which can be chosen freely subject to the economic budget constraint and the time constraint. Yet, in reality, it seems that "type of work" or "non-pecuniary job attributes" does matter a great deal in agents' labor supply decisions. The framework proposed in van-Ophem, Hartog and Vijverberg (1993), Aaberge, Dagsvik and Strøm (1995), and Dagsvik and Strøm (1998) acknowledge these qualitative aspects of jobs. In this paper, we adopt a similar approach to the one in Aaberge et al. (1995) by modeling labor supply behavior as a discrete choice problem, where the alternatives are viewed as "job packages". These job packages are characterized by specific attributes such as hours of work, wage rates and other non-pecuniary variables. Consequently, the set of feasible jobs will in this setting matter for whether or not the agent will be employed.

While many authors in previous studies have investigated the nature of the constraint on the labor market, their efforts have been directed towards restrictions on working hours. Ilmakunnas and Pudney (1990) formulate a labor supply model which is a mixture of logit-type models across unobservable choice sets, where the choice sets are equal to, or subsets of, the set consisting of the alternatives "part-time", "full-time" and "non-participation". Dickens and Lundberg (1993) and Tummers and Woittiez (1991) have proposed a model in which an individual receives a random number of job offers, each associated with weekly hours drawn from a discrete distribution common to all individuals. Connolly (1997) estimates a model for the probability of being employed which equals the probability of labor force participation times the probability of finding employment given participation.

In this paper we only consider, similarly to Connolly (1997), the modeling of the binary decision problem of choosing between "working" and "not working". More precisely, we assume that the researcher only observe the binary choices. However, in contrast to Connolly (1997) we take into account that the individual agent has preferences over hours of work as well as over (unobservable) job attributes, and that the set of feasible jobs is typically unobserved by the researcher. The size of the choice set of jobs might depend on variables such as regional unemployment rate, and individual characteristics such as education level and work experience. However, in our study, due to the lack of the data of involuntary unemployment, we assume that the choice set only depends on regional unemployment rate.

We estimate two versions of the model for "working" versus "not working" based on data from a sample of married women. In the first version we only use micro-data from independent cross-sections. In the second version we use panel data and allow preferences to be serially correlated. The motivation for introducing serially correlated random terms is that there may be taste persistence effects implying that the decision-makers' preference evaluation in successive periods will be correlated.

The paper is organized as follows. In Section 2, the basic model is developed. The empirical specification of the model is derived in Section 3. Data and estimation results are described in Section 4. In Section 4.2 we assume that the parameters are period-specific. In Section 4.3 we use panel data to estimate the model based on a method that takes into account serially correlated preferences. Different types of elasticities are calculated.

2. The model

In the simplest version of the static labor supply model, the individual's utility depends on his tastes and on the amount of consumer goods and hours of leisure time that he consumes per period. In this paper we shall extend this simple textbook model to include non-pecuniary attributes of jobs. The motivation is that attributes other than wage and working hours may have significant influence on the degree of an individual's "enjoyment" obtained from working. Examples of non-pecuniary attributes are the time spent on the way to workplace, the number of coffee breaks one can take at work, the facilities, the nature of the job, and the environment at the workplace. These attributes are unobservable to the researcher, but they sometimes play an important role in individuals' labor supply decision. Also we wish to take into account that pecuniary attributes such as wage rates vary across jobs.

The utility function individual *i* is assumed to have the structure

(1) $$\tilde{U}_i\left(C,h,\psi_i\right)=U_i\left(C,\sum_{k=1}^{m}\psi_{ik}h_k\right),$$

where C denotes consumption, h_k is annual hours of work associated with job k, $\psi_i=\left(\psi_{11},\psi_{12},...\right)$ is an unknown parameter vector which captures non-pecuniary aspects of job k as evaluated by individual i. The function $U_i(\cdot)$ is assumed to be quasiconcave and twice differentiable and is allowed to be individual specific.

Note that with this structure of the utility function the different jobs are perfect substitutes, that is, conditional on the consumption level, job *k* yields the same utility as job *j* as long as the adjusted working hours, $\psi_{ik} h_k$ and $\psi_{ij} h_j$, are equal, cf. Hanemann (1984). In other words, given the consumption level C and the non-pecuniary aspects of the jobs, ψ_{ik} and ψ_{ij}, if job j is less attractive than job k, by reducing the working hour of job j, h_j, one can make it as attractive as or even more attractive than job k.

The budget constraint equals

(2)
$$C = \sum_{k \in B_i} W_{ik} h_k + I \,,$$

and the time constraint equals

(3)
$$\sum_{k \in B_i} h_k + L = M \,,$$

where B_i is the set of jobs that are feasible to individual i, I is non-labor income, W_{ik} is the offered wage rate of job k to individual i, L denotes annual leisure hours, and M is total available time per year. Note that we have introduced B_i to represent the constraint imposed on the individual's opportunities from the demand side. The set B_i is unobservable to the researcher. Note also that, as mentioned above, we allow wage rates to be job-specific in this setup. Only the wage rate associated with the *chosen* job is observed.

Wages of other family members are included in I. The earnings of the male partner is treated as an exogenous variable. Furthermore, the budget constraint is assumed to be linear in hours of work. In other words, taxes are not taken into account. Now let $\psi_{ik} h_k = x_{ik}$, which implies that the corresponding utility maximization problem can be written as

(4)
$$max U_i \left(C, \sum_{k \in B_i} x_{ik} \right)$$

subject to

(5)
$$C = \sum_{k \in B_i} x_{ik} W_{ik}^* + I$$

and

(6)
$$\sum_{k \in B_i} \frac{x_{ik}}{\psi_{ik}} + L = M$$

where $W_{ik}^* = W_{ik} / \psi_{ik}$. We shall call $\{W_{ik}^*, k \in B_i\}$ the modified wage rates of individual *i*. From this it follows that the agent would like to supply labor if

(7) $$\max_{k \in B_i} W_{ik}^* > -\frac{\partial_2 U_i(I,0)}{\partial_1 U_i(I,0)} \equiv \tilde{W}_i$$

where $\tilde{W}_i$ is the reservation wage rate of individual *i*, and $\partial_j, j = 1,2,$ denotes the partial derivative with respect to component *j*. If the inequality holds, the maximization problem implies a "corner" solution where the individual selects the very job with the highest modified wage rate. This is due to the assumption of perfect substitution between jobs.

3. An empirical specification

In this section, we shall discuss an empirical specification based on the framework developed in Section 2. The decision problem discussed above yields a decision rule for whether or not to work expressed in terms of the modified wage rates, the set of feasible jobs and the reservation wage. Education and years of labor force experience are generally considered the main determinants of the offered wage rates. The reservation wage is assumed to depend on age, non-labor income and number of children. Those variables all arise from previous economic choices or chance events. Thus we assume that:

(8) $$\log W_{ij} = X_i \gamma + \eta_{ij} ,$$

and

(9) $$\log \tilde{W}_i = Z_i \beta + \varepsilon_{i0} ,$$

where X_i is a vector of individual characteristics such as education, experience and experience squared, Z_i is a vector consisting of age, age squared, non-labor income and number of children of age 0-6 and of age 7-17, η_{ij} and ε_{i0} are random terms that accounts for the unobservable effects, and γ and β are vectors of parameters of appropriate dimension. Furthermore we assume that we can write

(10) $$\log \psi_{ij} = b + \varepsilon_{ij}^*$$

where b is a constant and ε_{ij}^* are i.i.d. random terms. As a result, W_i^* can be written as,

(11) $$\log W_{ij}^* = \log W_{ij} - \log \psi_{ij} = \log W_{ij} + \varepsilon_{ij}^* = X_i \gamma + \eta_{ij} + \varepsilon_{ij}^* = X_i \gamma + \varepsilon_{ij}$$

where $\varepsilon_{ij} = \eta_{ij} + \varepsilon^*_{ij}$ and *b* is absorbed in the intercept of *Xγ*. This represents no loss in generality. We assume next that $\theta\varepsilon_{ij}$, $j = 0,1,...,n,$ are i.i. standard extreme value (type III) distributed for some $\theta > 0$ which means that

(12) $$P\left(\theta\varepsilon_{ij} \leq y \middle| X_i\right) = \exp\left(-e^{-y}\right).$$

This assumption can be justified from Luce Choice Axiom "Independence from Irrelevant Alternatives" (IIA), see for example Ben-Akiva and Lerman (1985).

Recall that IIA can be interpreted as follows: Suppose that the agent faces the choice set *B*. He then first selects a subset *A* from *B* where *A* contains alternatives which are preferable to the alternatives in *B\A*, and subsequently makes his choice from the subset *A*. So far there is no essential loss of generality, since usually it is always possible to think of the decision process in this manner. The crucial property of IIA is that, on average, the choice from *A* in the latter stage does not depend on alternatives outside *A*. In other words, the alternatives discarded in the former stage has been completely "forgotten" by the agent and they are no longer relevant when the agent makes his second stage decision. One thing worth mentioning is that, by using the phrase "on average", we mean that it is not necessary for IIA to hold in every experiment. Only when the choice experiment has been replicated a large number of times, or equivalently, when the experiment is conducted on a large sample of "identical agents" (agents with identically distributed tastes), IIA should hold on average. In this sense, we may think of IIA as an assumption of "probabilistic rationality".

When the error term in the utility function has a standard type III extreme value distribution, given in (12), the random utility model become consistent with the IIA property, and this will in turn give us a simple structure of the choice probabilities.

The parameter, *θ*, introduced above can be interpreted as

(13) $$\theta^2 = \frac{\pi^2}{6 Var\varepsilon_{ij}}.$$

The error term, η_{ij}, is likely to be correlated with ε^*_{ij}, because the unobserved variables that affect the wage rates $\{W_{ij}\}$ are likely to correlate with unobservable taste shifters associated with non-pecuniary aspects of the jobs. This is actually one of the reasons why we introduced an instrument variable equation like eq. (8) in the first place. The other reason is that the (potential) wage rates for those who do not work and for jobs that are not chosen are unobserved.

Following the assumption about the distribution of $\{\varepsilon_{ij}\}$ the probability of choosing job j from choice set B_i can be expressed as

(14) $$P_{ij} = P\left(W_{ij}^* = \max\left(\tilde{W}_i, \max_{k\in B_i} W_{ik}^*\right)\right) = \frac{e^{\theta X_i \gamma}}{e^{\theta Z_i \beta} + \sum_{k\in B_i} e^{\theta X_i \gamma}} = \frac{e^{\theta X_i \gamma}}{e^{\theta Z_i \beta} + n_i e^{\theta X_i \gamma}},$$

where n_i is the number of jobs available to the individual, i.e. the number of jobs in B_i. The probability of working therefore equals

(15) $$P_{i2} = \sum_{k\in B_i} P_{ik} = \frac{n_i e^{\theta X_i \gamma}}{e^{\theta Z_i \beta} + n_i e^{\theta X_i \gamma}} = \frac{1}{1+\exp\left(\theta Z_i \beta - \theta X_i \gamma - \log n_i\right)}.$$

This type of approach intended to take into account latent choice sets is similar to the approach described in Ben-Akiva and Lerman (1985), p.p. 253-261. We assume that n_i depends only on regional unemployment rates. This assumption may seem restrictive, for one could argue that n_i should also depend on education level and work experience of the individual. Nevertheless, the present assumption is adopted to enable us to identify all the parameters in the model. If individual observations on unemployment were available we would be able to identify n_i as a function of for example education and experience, similarly to Connolly (1997). The variable length of work experience is defined as age minus years of education minus six. By using (15) we are able to estimate the reduced form probability of working. However, we need to estimate the offered wage equation to obtain identification of $\theta\gamma$ and $\theta\beta$.

It is well known that a sample consisting only of working individuals is not necessarily a random sample, and this may lead to biased estimates if simple methods such as ordinary least squares (OLS) are applied directly. Still, the problem can be solved by using a version of Heckman's (1979) two-stage method. Let J denote the chosen job within B_i. We assume that

(16) $$E\left(\eta_{ij} \middle| \varepsilon_{ij}\right) = \zeta\left(\theta\varepsilon_{ij} - 0.5772\right),$$

where ζ is a constant. Under this assumption one can show that

(17) $$\log W_{iJ} = X\gamma - \zeta \log P_{i2} + \eta_{iJ}^*,$$

where η_{iJ}^* is a random term with the property

$$E\left(\eta_{iJ}^* \mid W_{iJ}^* > \tilde{W}_i\right) = 0.$$

The formal derivation of eq. (17) is relegated to Appendix 1.

4. Estimation method and results

4.1. Data description

The data are obtained from the Income Distribution Survey in 1989, and 1992-1994, to which additional information about working hours and regional unemployment rate is appended. All the data sets were collected and prepared by Statistics Norway and all the participants are Norwegian citizens.

The Income Distribution Survey consists mainly of tax return data collected from taxation authorities. Certain criteria have been adopted when the sample was selected. Only married women in the age between 25 and 64 are selected (husbands' information is also included). Couples for which one or both have entrepreneurial income in excess of wage income are excluded.

We have converted the variable "length of education" from category values to numbers. Subtracting hours worked as self-employed from total working hours, which is the product of number of working weeks and average working hours per week, constitutes the variable "working hours". The (marginal) wage rate is measured as wage income divided by working hours.

For unknown reasons, information about working hours is missing for most of and the whole sample in 1990 and 1991, respectively. Furthermore, for survey 1989, 1992-1994, the rate of non-response was almost 50 per cent. Therefore, the data set is an unbalance panel. During the whole period (i.e. 1989, 1990, 1992-1994), there are in total 1057 different women participating in the survey, however, due to missing values to one or more categories for some of them, we end up with only 999 participants. The average years of numbers of participation is 2.83. In this paper, women who worked for less than 100 hours per year are registered under category "not working". This differs from the definition applied in Dagsvik and Strøm (1995), in which they use 60 hours instead. The reason is that very few persons in our sample work less than 60 hours per year, and, accordingly, if one should adopt the definition of Dagsvik and Strøm (1995), the average probability of working would be close to one. Below we report some summary statistics for the sample.

Table 1. Summary statistics for the whole sample during 1989, 1992-1994*)

	1989		1992		1993		1994	
	The whole sample	The sub-sample: those who work	The whole sample	The sub-sample: those who work	The whole sample	The sub-sample: those who work	The whole sample	The sub-sample: those who work
No. of observations	412	363 (88%)	551	490 (89%)	523	476 (91%)	521	462 (89%)
Non-labour income	190 346 (81 831)	186 820 (73 489)	225 708 (91 113)	223 113 (88 752)	236 006 (96 702)	232 684 (94 687)	252 034 (94 288)	248 019 (90 949)
No. of children of age 0-6	0.42 (0.74)	0.37 (0.70)	0.31 (0.66)	0.26 (0.60)	0.30 (0.67)	0.26 (0.62)	0.26 (0.63)	0.23 (0.58)
No. of children of age 7-17	0.75 (0.87)	0.75 (0.86)	0.72 (0.89)	0.71 (0.87)	0.74 (0.88)	0.71 (0.85)	0.67 (0.83)	0.66 (0.81)
Age	41 (7.93)	41 (7.91)	44 (7.46)	44 (7.29)	44 (7.33)	44 (7.14)	44 (7.28)	45 (6.93)
Length of education	11 (1.96)	11 (1.96)	11 (2.00)	11 (2.03)	11 (1.99)	11 (2.00)	11 (2.03)	11 (2.05)
Length of experience	24 (8.56)	24 (8.54)	27 (8.05)	27 (7.90)	27 (7.90)	27 (7.68)	28 (7.84)	28 (7.46)
Regional unemploy-ment	0.019 (0.008)	0.019 (0.008)	0.026 (0.007)	0.026 (0.007)	0.027 (0.007)	0.027 (0.007)	0.025 (0.007)	0.025 (0.007)
Wage rate per hour		92.73 (34.23)		106.35 (42.31)		107.34 (41.01)		112.79 (43.93)

*) Except for on the line of "no. of observation" where per cents of women who work are given in the parentheses, standard deviations are given in the parentheses.

Each year about 90 per cent of the sample work. The figures of average non-labor income and average no. of children of age 0-6 and 7-17 in the sub-sample are slightly smaller than those in the whole sample, which seems to be in accord with intuition. The average wage rate per hour increases slightly over time.

4.2. Estimation results

In the first version of the model, we allow parameters to be period-specific. We can therefore estimate the model as if the cross-sections of each year were independent. To use the method mentioned above to identify all the parameters in (15) will lead to a three-stage estimation procedure: First, we estimate the reduced form probability of working. In the second stage we estimate the wage equation and due to eq. (17) we use the logarithm of the (estimated) probability of working as an additional regressor to control for selection bias. Finally, the structural form of probability of working is estimated.

The results from stage one are displayed in Table 2.

Table 2. Estimation results for the reduced form employment probability*)

Variable / Year	Constant	log of non-labor income	No. of children of age 0-6	No. of children of age 7-17	Regional un-employ-ment rate	Age	Edu-cation	Age squared × 10^{-2}	2 Age times (Edu-cation - Edu-cation squared)
1989	-16.25 (-1.64)	0.79 (2.17)	0.83 (2.96)	0.16 (0.72)	0.88 (0.04)	0.26 (0.75)	0.20 (0.42)	-0.093 (-0.34)	-0.87 (-0.97)
1992	-4.97 (-0.53)	0.80 (2.20)	1.29 (5.21)	0.21 (1.24)	21.62 (0.92)	-0.11 (-0.38)	-0.56 (-1.06)	0.14 (0.62)	0.10 (0.11)
1993	-16.60 (-1.50)	1.25 (3.07)	1.37 (4.53)	0.97 (4.45)	-23.50 (-1.03)	-0.14 (-0.40)	0.12 (0.23)	0.44 (1.64)	-0.89 (-0.93)
1994	-17.85 (-1.99)	1.61 (3.53)	1.13 (4.15)	0.61 (2.97)	-11.49 (-0.48)	-0.23 (-0.96)	0.07 (0.23)	0.48 (2.18)	-0.69 (-1.31)

*) t-statistics are given in the parentheses.

Notice that the estimates of the coefficients of non-labor income and no. of children at the age between 0 and 6 are significant in all periods, and have the expected sign. In addition, they do not vary much over time. The estimated values for the variable, no. of children between 7 and 17, though not significant in all the surveyed year, have nevertheless also an expected sign. The positive coefficients for those variables indicate that an increase in any one of them decreases a married woman's probability of choosing to work.

Other coefficients associated with variables such as age and education level, are not significantly different from zero.

The results from the second stage are presented in Table 3.

Table 3. Estimation results for the wage equation

Year \ Parameter	Intercept	Education	Experience	Experience squared × 10^{-2}	log of working probability
1989	4.05 (18.82)	0.033 (3.20)	0.0085 (0.80)	-0.014 (-0.70)	-0.06 (-0.25)
1992	3.78 (17.76)	0.029 (3.21)	0.004 (0.35)	$-4.2 \cdot 10^{-4}$ (-1.24)	0.44 (2.16)
1993	4.59 (19.64)	0.038 (4.24)	0.007 (0.50)	-0.02 (-0.81)	-0.46 (-2.01)
1994	4.18 (17.94)	0.036 (3.90)	0.002 (0.12)	0.004 (0.14)	0.02 (0.01)

*) t-statistics are given in the parentheses.

It seems that education level is the main determinant of the offered wage rate. Experience does not have significant influence on offered wage, and the selection effect is weak. Moreover, there is no evident trend in the intercept and the coefficient associated with education seems to be rather stable over time.

We are now able to estimate the structural form probability of working. The results are displayed in Table 4.

Table 4. Estimation results for the structural employment probability

Year \ Variable	Constant	log of non-labor income	No. of children at the age 0-6	No. of children at the age 7-17	Regional unemploy -ment rate	Age	Age squared × 10^{-2}	log wage rate
1989	-10.57 (-1.68)	0.70 (2.02)	0.86 (3.10)	0.13 (0.61)	2.05 (0.09)	0.16 (0.66)	-0.16 (-0.61)	8.42 (2.47)
1992	-6.43 (-0.95)	0.80 (2.21)	1.28 (5.26)	0.22 (1.28)	18.42 (0.93)	$0.8 \cdot 10^{-3}$ ($3.7 \cdot 10^{-3}$)	0.12 (0.56)	19.85 (4.34)
1993	-12.97 (-1.67)	1.29 (3.12)	1.44 (4.93)	0.98 (4.49)	-35.39 (-1.73)	-0.19 (-0.76)	0.26 (0.96)	10.01 (3.67)
1994	-10.99 (-1.39)	1.55 (3.40)	1.18 (3.40)	0.60 (2.87)	-3.03 (-0.13)	-0.39 (-1.83)	0.55 (2.52)	11.21 (3.77)

*) t-statistics are given in the parentheses.

From Table 4 we see that the coefficient of the logarithm of the wage rate is significant, but vary much from year to year. However, if we apply 95 per cent confidence intervals based on the estimated standard deviations in each period we realize that none of these estimates are significantly different. The estimated parameters for non-labor income and no. of children between 0 and 6 have the right signs and are significant. In addition, the estimates are quite similar to those obtained in stage one. Other estimates are unfortunately not significant.

4.3. Accounting for serial correlation

In the analysis above we have estimated the model with period specific parameters. We shall now assume that the parameters in the model are constant over time. This implies that we need to take into account the panel structure of the data since there may be serial correlation in the agents' labor market choices over time. In fact, we shall assume that the choice model is a Markov chain, which means that the model can be completely represented through transition probabilities or, alternatively, transition intensities. However, we do not have the ambition to formulate a full structural model of the transition probabilities (or transition intensities). As above, we only postulate a structural model for the probability of working versus not working, but we wish to impose a structure of the transition intensities which is consistent with the (given) structural specification of the probability of working.

For simplicity we drop the indexation of the individual in this section. Let $\{q_{jk}(t)\}$ denote the transition intensities of the Markov chain with two states ("working" and "not working") in continuous time, $j,k=1,2$. This means that $q_{jk}(t)\Delta t$ is (approximately) the probability of switching to state *k* in the time interval $(t,t+\Delta t)$ given that state *i* was occupied at time *t*. For example, $q_{12}(t)\Delta t$ is the probability that an agent who does not work at time *t* will change to working status within $(t,t+\Delta t)$.

Suppose for the moment that the transition intensities are constant over time. Then it follows that the probabilities of being in state 1 and 2 equal

(18) $$P_1=\frac{q_{21}}{q_{12}+q_{21}}, \quad P_2=\frac{q_{12}}{q_{12}+q_{21}}.$$

Hence, we can write

(19) $$q_{21}=\varphi P_1 \text{ and } q_{12}=\varphi P_2,$$

where

(20) $$\varphi=q_{12}+q_{21}.$$

Thus, for any set of transition intensities q_{12} and q_{21}, we can relate these to the probabilities of being in the respective states, P_1 and P_2, through the construction (19). Specifically, this means that when P_j, $j=1,2,$ is given then we can represent the law of the Markov chain by these probabilities and $\varphi=q_{12}+q_{21}$.

Furthermore, it follows from the probability theory of Markov chains that the corresponding transition probabilities for transitions from time s to time t, are equal to

(21) $$Q_{21}(s,t) = \left(1 - e^{-(t-s)\varphi}\right) P_1,$$

(22) $$Q_{12}(s,t) = \left(1 - e^{-(t-s)\varphi}\right) P_2$$

and

(23) $$Q_{jj}(s,t) = P_j + e^{-(1-s)\varphi}\left(1 - P_j\right),$$

for $j = 1,2$. We realize that when $\varphi = \infty$, there is no serial dependence, i.e. the preferences are not serially correlated because in this case $Q_{jk}(s,t) = P_k(t)$. In contrast, when $\varphi = 0$ then $Q_{jk}(s,t) = 0$ for $j \neq k$, which means that the unobservables of the utility function is perfectly correlated over time. Thus $\varphi \geq 0$ can be interpreted as a measure of taste persistence of the agents preferences. The parameter φ can also be given a more direct and explicit interpretation associated with the autocorrelation function of the choice process, suitably defined. To realize this, let $Y_j(t)$ be equal to one if the individual is in state *j* at time *t*, and zero otherwise. Then we have that

(24) $$\begin{aligned} Cov\left(Y_j(t), Y_j(s)\right) &= E\,Y_j(t)Y_j(s) - E\,Y_j(t)\,E\,Y_j(s) \\ &= Q_{jj}(s,t)P_j - P_j^2 = P_j^2 + e^{-(t-s)\varphi}\left(1 - P_j\right)P_j - P_j^2 \\ &= e^{-(1-s)\varphi} P_j \left(1 - P_j\right). \end{aligned}$$

But since

(25) $$Var\,Y_j(t) = Var\,Y_j(s) = P_j\left(1 - P_j\right)$$

it follows that

(26) $$Corr\left(Y_j(t), Y_j(s)\right) = e^{-(t-s)\varphi}.$$

We have thus demonstrated that

$$e^{-(t-s)\varphi}$$

can be interpreted as the autocorrelation function of the process $\{Y_j(t), t>0\}$, $j=1,2$. In the general case with nonstationary transition intensities, that are constant within years, it can be shown similarly that

(27) $$Q_{jk}\left(t-1,t\right)=\left(1-e^{-\varphi(t)}\right)P_k\left(t\right)$$

for $j \neq k$ and

(28) $$Q_{ii}\left(t-1,t\right)=P_j\left(t\right)+e^{-\varphi(t)}\left(1-P_j(t)\right)$$

where

(29) $$\varphi(t)=q_{12}\left(t\right)+q_{21}\left(t\right).$$

In the estimation below we assume that $\varphi(t)$ is approximately constant over time so that we can set $\varphi(t)=\varphi$. Note that this is an approximation because in general φ may depend on the explanatory variables of the model. Recall moreover that we assume that $P_j(t)$ has the same structure as in eq. (15) but now with time constant parameters. The estimation results are displayed in Table 5.

Table 5. Estimation results with serial correlation

	Constant	log of non-labor income	No. of children at the age 0-6	No. of children at the age 7-17	Regional unemploy-ment rate	Age	Age squared $\times 10^{-2}$	log wage rate	φ
Parameter	-6.76	0.96	0.97	0.35	7.86	-0.19	0.28	8.88	0.23
t-statistic	-1.71	4.23	5.26	3.02	0.58	-1.47	2.02	4.81	6.69

The small value of φ indicates strong taste persistence, as we expected. Specifically, the estimate of φ implies that the correlation between $Y_j(t)$ and $Y_j(t-1)$ equals 0.79. Other findings are also not surprising. As in our previous time-independent model, non-labor income and number of children and log offered wage rate are the variables which have significant influence on the labor supply decision.

As a measure of the goodness of fit, we have calculated the measure proposed by McFadden. The measure is defined as

(29) $$\rho^2=1-\frac{L\left(\hat{\beta}\right)}{L\left(0\right)}$$

where $L(\hat{\beta})$ is the value of the log likelihood function at the estimated parameters and $L(0)$ is its value when φ is set equal to infinity and all the remaining parameters are set equal to zero. The value of ρ^2 is found to be 0.69. The loglikelihood $L(\hat{\beta})$ equals -450.

5. Policy simulations

It is often of interest to know the impact on the probability of working as a result of the marginal change in the explanatory variables. Below we will report elasticities with respect to non-labor income and mean wage rate for selected values of explanatory variables.

Table 6. Simulated marginal effects of changes in expected wage, non-labor income, and number of children

Wage rate	Non-labor income	No. of children at the age 0-6	No. of children at the age 7-17	Regional unemploy-ment rate	Age	Prob. of working	Wage elas.	Non-labor income elas.	Change in the probability of working due to a marginal increase in no. of small children	Change in the probability of working due to a marginal increase in no. of large children
85.37	100000	0	0	0.02	30	0.938	0.554	-0.060	-0.087	-0.024
99.19	100000	0	0	0.02	30	0.983	0.153	-0.017	-0.027	-0.007
115.24	100000	0	0	0.02	30	0.995	0.041	-0.004	-0.007	-0.002
85.37	300000	0	0	0.02	30	0.840	1.420	-0.153	-0.174	-0.053
99.19	300000	0	0	0.02	30	0.952	0.425	-0.046	-0.069	-0.019
115.24	300000	0	0	0.02	30	0.987	0.116	-0.013	-0.021	-0.005
85.37	300000	1	0	0.02	30	0.666	2.963	-0.320	-0.235	-0.082
99.19	300000	1	0	0.02	30	0.883	1.037	-0.112	-0.141	-0.041
115.24	300000	1	0	0.02	30	0.966	0.299	-0.032	-0.050	-0.014
85.37	500000	0	0	0.02	30	0.763	2.104	-0.227	-0.213	-0.069
85.37	300000	0	1	0.02	30	0.787	1.892	-0.204	-0.203	-0.065
99.19	300000	0	1	0.02	30	0.933	0.592	-0.064	-0.092	-0.026
115.24	300000	0	1	0.02	30	0.981	0.164	-0.018	-0.029	-0.008

An increase in non-labor income, provided other things equal, decreases the probability of working, and give rise to non-labor income elasticities (in absolute value; in this section we shall always compare figures in their absolute value) and the others. Note that when non-labor income equals 100 000 the wage elasticities are lower than when non-labor income is higher than 100 000. The reason for this is that in this case the probability of working is close to one which means that most women are working.

An increase in either the number of children at the age between 0 and 6 or the number of children at the age between 7 and 17 will reduce the probability of working and boost the marginal influence of all the variables, but with different magnitude. An alternation in the number of small children has larger effect on the woman's decision, which is in accordance with our estimation results in Section 4, namely, that the estimated parameter for number of small children is greater than that of number of big children. This is also the case irrespective of the level of non-labor income.

Sometimes aggregate elasticities of for example wage and non-labor income may be more useful to the policy maker than the average individual elasticities. As regard to aggregate elasticity, I shall write down explicitly the formulas used to avoid misunderstanding, since, albeit widely used, the concept is not commonly known as elasticity. It is defined as follows:

$$(30)\quad Aggregate\ El_{x_i}^{P_{2i}} = \frac{\sum_i \left(P_{2i}\left(\left(1+r\right)x_i\right) - P_{2i}\left(x_i\right)\right)}{\sum_i P_{2i}\left(x_i\right)} \cdot \frac{x_i}{\left(1+r\right)x_i - x_i} = \frac{\sum_i \left(P_{2i}\left(\left(1+r\right)x_i\right) - P_{2i}\left(x_i\right)\right)}{r\sum_i P_{2i}\left(x_i\right)}$$

where x_i denotes the relevant variable such as, for example, non-labor income, etc., r is a percentage figure, like 10 per cent, for example, $P_{i2}\left(\left(1+r\right)x_i\right)$ is the probability of working after x_i has been raised by r per cent given other things equal, and $P_{i2}\left(x_i\right)$ is the probability of working with original value of variable x_i. In general, the aggregate elasticity will differ from the average elasticity

$$\frac{1}{N}\sum_{i=1}^{N}\frac{\partial \log P_{i2}}{\partial \log x_i}.$$

Only when $P_{i2}, i = 1,2,...,N,$ are identical across the sample, will the two figures become equal. We therefore calculate both the aggregate and the average elasticities for selected variables. The results are displayed in Table 7.

Table 7. Average and aggregate effects

Year	Average wage elasticities	Average non-labor income elasticities	Average percentage change in the probability of working if every woman has an additional child at the age 0-6	Average percentage change in the probability of working if every woman has an additional child at the age 7-17	Aggregate wage elasticities	Aggregate non-labor income elasticities	Aggregate percentage change in labor supply if every woman has an additional child at the age 0-6	Aggregate percentage change in labor supply if every woman has an additional child at the age 7-17
1989	0.97	-0.10	-0.10	-0.03	0.63	-0.09	-0.13	-0.04
1992	0.95	-0.10	-0.08	-0.03	0.62	-0.09	-0.13	-0.04
1993	0.98	-0.11	-0.08	-0.03	0.64	-0.10	-0.13	-0.04
1994	1.02	-0.11	-0.07	-0.03	0.67	-0.10	-0.14	-0.04

In the above table, r is chosen to be 10 per cent when we calculate the figures in column 5 and 6. Those figures vary little across years.

6. Conclusions

In this paper, we have accommodated the notion of latent sets of job opportunities and job specific non-pecuniary attributes in the specification of the model of labor supply. We have focused only on the binary choice, that is, "working" and "not working". We have applied a particular method to account for serial correlation when estimating the model from panel data. The estimates and the calculated elasticities appear to be plausible.

However, our specification of the parameters associated with the latent size of the choice set of jobs is somehow unsatisfactory due to the lack of data on unemployment.

References

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

Ben-Akiva, M., and S. Lerman (1985) : *Discrete Choice Analysis: Theory and Application to Travel Demand.* MIT Press, Cambridge, Massachusetts.

Connolly, S. (1997): A Model of Female Labour Supply in which Supply is Dependent upon the Chances of finding a Job. *Applied Economics*, **29**, 1379-1386.

Dagsvik, J.K. (1999): Probabilistic Models for Qualitative Choice Behavior. An Introduction. Document, Statistics Norway.

Dagsvik, J.K. and S. Strøm (1995): Labor Supply with Non-Convex Budget Sets, Hours Restrictions and Nonpecuniary Attributes. Mimeo, Statistics Norway.

Dagsvik, J.K., D.G. Wetterwald and R. Aaberge (1996): Potential Demand for Alternative Fuel Vehicles. *Discussion Papers*, no. **165**, Statistics Norway.

Dickens, W., Lundberg, S. (1993): Hours Restrictions and Labor Supply. *International Economic Review*, **34**, 169-191.

Hanemann, W.M. (1984): Discrete/Continuous Choice of Consumer Demand. *Econometrica*, **52**, 541-561.

Heckman, J.J. (1979): Sample Seclection Bias as a Specification Error. *Econometrica*, **47**, 153-161.

Ilmakunnas, S., Pudney, S. (1990): A Model of Female Labour Supply in the Presence of Hours Restrictions. *Journal of Public Economics*, **41**, 183-210.

Killingsworth, M.R., Heckman, J.J. (1986): Female Labor Supply: A Survey: *Handbook of Labor Economics*, Vol. I, Elsevier Science Publishers BV, New York.

Strauss, D. (1979): Some Results on Random Utility Models. *Journal of Mathematical Psychology*, **20**, 35-52.

Tummers, P.M., Woittiez, I. (1991): A simultaneous Wage and Labor Supply Model with Hours Restrictions. *Journal of Human Resources*, **26**, 393-423.

van-Ophem, H., Hartog, J., Vijverberg, W.M. (1993): Job Complexity and Wages. *International Economic Review*, **34** (4), Nov 1993, 853-72.

Derivation of the selection effect in the wage equation

The derivation of eq. (17) is given in Dagsvik (1999). For the readers' convenience I will go through the basic argument below. First we need the following lemma found in Strauss (1979).

Lemma 1

Suppose $U_j = \mathrm{v}_j + \varepsilon_j$, *where* ε_j, $j = 1,2,...,$ *are i.i.d. with (standard) cumulative distribution function* $\exp\left(-e^{-y}\right)$. *Then*

$$P\left(max_k U_k \le y \mid U_j = max_k U_k\right) = P\left(U_j \le y \mid U_j = max_k U_k\right) = P\left(max_k U_k \le y\right).$$

We assumed in Section 3 that $\theta\varepsilon_{ij}$, $j = 0,1,...,n,$ are i.i. standard extreme value distributed. Recall that J denotes the chosen job, i.e. $W_{iJ}^* = \max_{k \in B_i} W_{ik}^*$. Under the preceding assumptions, the distribution of W_{iJ}^* can be readily derived. Specifically,

$$\begin{aligned} P\left(\theta \log W_{iJ}^* \le y\right) &= P\left(\bigcap_{k \in B_i}\left(\theta \log W_{ik}^* \le y\right)\right) \\ &= \prod_{k \in B_i} P\left(\theta \log W_{ik}^* \le y\right) = \prod_{k \in B_i} \exp\left(-e^{\theta X_i \gamma - y}\right) \\ &= \exp\left(-n_i e^{\theta X_i \gamma - y}\right) = \exp\left(-\exp\left(\log n_i + \theta X_i \gamma - y\right)\right). \end{aligned} \tag{A.1}$$

Similarly, it follows that

$$P\left(\max\left(\log W_{iJ}^*, \log \tilde{W}_i\right) \le y\right) = \exp\left[-\exp\left(\log\left(n_i\, e^{\theta X_i \gamma} + e^{\theta Z_i \beta}\right) - y\right)\right]. \tag{A.2}$$

From Lemma 1 we obtain that

$$P\left(\theta \log W_{ij}^* \le y \middle| W_{iJ}^* > \tilde{W}_i\right) = P\left(\max\left(\theta \log W_{iJ}^*, \theta \log \tilde{W}_i\right) \le y\right) \tag{A.3}$$

which implies that

$$E\left(\theta \log W_{ij}^* \middle| W_{iJ}^* > \tilde{W}_i\right) = \theta E \max\left(\log W_{iJ}^*, \log \tilde{W}_i\right) = 0.5772 + \log\left(n_i\, e^{\theta X_i \gamma} + e^{\theta Z_i \beta}\right). \tag{A.4}$$

Similarly (A.1) yields

(A.5) $$E\theta \log W_{iJ}^{*} = 0.5772 + \log n_i + \theta X_i \gamma .$$

Hence, by using (16), (A.2) to (A.5) we get

(A.6)
$$\begin{aligned}
&E\left(\eta_{iJ} \left| W_{iJ}^{*} > \tilde{W}_i \right.\right) = \zeta E\left(\left(\theta \varepsilon_{iJ} - 0.5772\right) \left| W_{iJ}^{*} > \tilde{W}_i \right.\right) \\
&= \zeta E\left(\theta \log W_{iJ} - \theta E \log W_{iJ} \left| W_{iJ}^{*} > \tilde{W}_i \right.\right) \\
&= \zeta E\left(\theta \log W_{iJ} \left| W_{iJ}^{*} > \tilde{W}_i \right.\right) - \zeta \theta E \log W_{iJ} \\
&= \zeta \theta E \max\left(\log W_{iJ}, \log \tilde{W}_i\right) - \zeta \theta E \log W_{iJ} \\
&= \zeta \log\left(n_i e^{\theta X_i \gamma} + e^{\theta Z_i \beta}\right) - \zeta \left(\log n_i + \theta X_i \gamma\right) \\
&= \zeta \log\left(\frac{n_i e^{\theta X_i \gamma} + e^{\theta Z_i \beta}}{n_i e^{\theta X_i \gamma}}\right) = \zeta \log\left(\frac{1}{P_{i2}}\right) = -\zeta \log P_{i2} .
\end{aligned}$$

Consequently, we can write

(A.7) $$W_{iJ} = Z_i \beta - \zeta \log P_{i2} + \eta_{iJ}^{*}$$

where

$$E\left(\eta_{iJ}^{*} \left| W_{iJ} > \tilde{W}_i \right.\right) = 0.$$

Thus one can estimate eq. (A.7) consistently from the subsample of working individuals by adding an estimate of P_{i2} as an additional regressor.