Discussion Papers No. 550, July 2008
Statistics Norway, Research Department

John K. Dagsvik and Zhiyang Jia

An Alternative Approach to Labor Supply Modeling
Emphasizing Job-type as Choice Variable

Abstract:
Traditional labor supply analysis is based on the assumption that workers only have preferences over consumption and hours of work, and are able to choose consumption and hours freely within the budget constraint. Recently, various discrete choice versions of the traditional approach (with discrete hours) have become popular, but the basic assumption above is still maintained. Neither of these two approaches allows for agents'preferences over qualitative job-specific or choice restrictions facing the agents in the labor market in terms of restricted choice sets of job opportunities.

In this paper we argue for an alternative modeling framework that differs from the standard models of labor supply in that the notion of job choice is fundamental. Specifically, the worker is assumed to have preferences over a latent worker-specific choice set of jobs from which he chooses his preferred job. A job is characterized with fixed (job-specific) working hours, wage rate and non-pecuniary attributes. As a result, observed hours of work and wage rate are interpreted as the job-specific (fixed) hours of work and wage rate associated with the chosen job.

The discussion in this paper focuses on interpretation of different versions and extensions of the alternative framework, theoretical and practical advantages, and how this approach relates to familiar existing approaches in the literature.

Keywords: Labor supply, non-pecuniary job attributes, non-convex budget sets, latent choice sets, random utility models.

JEL classification: J22, C51

Acknowledgement: We thank Jørgen Aasness, Thor Olav Thoresen, Terje Skjerpen and Ådne Cappelen for valuable comments.

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

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

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

For printed Discussion Papers contact:

Statistics Norway
Sales- and subscription service
NO-2225 Kongsvinger

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

1. Introduction

Behavioral labor supply analysis has been an important field of research in the last few decades. Although a variety of sophisticated approaches have been developed and applied for empirical analysis, there is less general agreement in the research community about the "preferred" modeling approach. In this paper, we discuss an alternative modeling strategy that differs from most existing ones, and where job type is the essential choice variable and we subsequently compare this strategy with traditional approaches. More precisely, the questions addressed in this paper are the following: (i) How should one formulate an empirical structural framework consistent with choice among job types ? (ii) What are the essential choice restrictions faced by the agents and how one may conveniently accommodate heterogeneity in these restrictions? (iii) How should functional form and distributional assumptions of preferences be selected and justified on theoretical grounds? (iv) How does our alternative modeling strategy relate to more traditional approaches?

In the basic version of the standard approach to labor supply analysis individual labor supply is viewed as a choice among feasible leisure and total consumption combinations within a convex budget set. See Blundell and MaCurdy (1999) for a review of different versions of the standard approach. An important generalization of this framework to accommodate non-convex budget sets was made by Blomquist, Burtless, Hausman and others (the Hausman approach), see for example Blomquist (1983, 1992), Burtless and Hausman (1978), Hausman (1985), Hausman and Ruud (1984), Heim and Meyer (2003), MaCurdy, Green and Paarsch (1990) and MaCurdy (1992). The non-convexity property of the budget set arises from typical tax and benefit systems in many countries. Although the Hausman approach was an important contribution at the time, experience has shown that this methodology is in general very difficult to apply except in special cases. Specifically, a recent study by Bloemen and Kapteyn (2008) demonstrates that even in the single agent case it is almost impossible to write down the true likelihood function of the empirical model given standard assumptions about unobservables, and considerable expertise and computer time is required to estimate this type of model. Thus, from a practical point of view it may seem questionable if the Hausman approach is at all a useful empirical strategy.

Recently, several versions of the discrete choice approach to labor supply modeling have been proposed. An advantage with the discrete choice approach is that it is much more practical to apply than the conventional continuous one. The reason is that whereas the continuous choice approach is based on conventional marginal criteria from which the supply function is derived, the discrete choice approach is based on global criteria and only requires specification of utility levels. As a result, it becomes easy to deal with nonlinear and nonconvex economic budget constraints within the discrete approach, unlike in the standard approach. Versions of discrete choice labor supply models have been estimated by Ilmakunnas and Pudney (1990), van Soest (1995) and van Soest, Das and Gong (2002). See also Blundell et al. (2000), Bingley and Walker (1997), Creedy and Kalb (2005), Duncan and Giles (1998), Duncan and Weeks (1998), Hoynes (1996), and Keane and Moffitt (1998), for different versions of the discrete choice approach.

A more fundamental theoretical issue is that the basic standard approach, including the modified one represented by the Hausman approach, ignore that an agent (supplier) in the labor market may have preferences over job attributes in addition to leisure and consumption combinations, and may face restrictions on his or her choice among job opportunities and hours of work. In this respect, a traditional discrete choice approach represents no essential departure from the standard approach because the only new feature introduced is that the set of feasible hours of work is finite. Thus, the problem of accommodating restrictions on hours of work remains within the usual discrete choice approach as well, and it is not easily dealt with. Apart from a few studies such as Blundell et al. (1987), Ilmakunnas and Pudney (1990), Kapteyn, Kooreman and van Soest (1990), Tummers and Woittiez (1990), van Soest, Woittiez, and Kapteyn (1990), Dickens and Lundberg (1993) and Bloemen (2000), the problem of rationing of jobs and restrictions on hours of work is typically ignored in most labor supply studies and the importance of the issue is generally undercommunicated within the research community. This continues to be the case in spite of the fact that the standard approach is unable to account for observed peaks at full-time and part-time hours of work found in most countries.

The alternative modeling approach discussed in this paper was, in the context of labor supply, initially proposed by Dagsvik and Strøm (1988) and Aaberge, Dagsvik and Strøm (1995) and further developed by Dagsvik and Strøm (1997, 2006). It is based on the work of Dagsvik (1994) and developed within a discrete choice framework.[1] In our approach, labor supply behavior is viewed as an outcome of agents' choices from a set of job 'packages'. Each package is characterized by an offered wage rate, offered hours of work and nonpecuniary (qualitative) attributes describing the nature of the job-specific tasks to be performed. For example, employment positions with the same tasks to be performed but with different working hours are viewed as different jobs. In some cases the researcher has access to observable qualitative attributes, which characterize the type of jobs and are relevant for agents' preference rankings. Unfortunately, most qualitative aspects of jobs cannot easily be represented by observable numerical attributes. In practice, it seems that researchers can at most

[1] Although the agent's choice problem is a discrete one (choice among feasible jobs), the distribution of the chosen hours of work may be continuous or discrete, depending on the assumptions about the distribution of the choice opportunities. See Dagsvik (1994) and Dagsvik and Strøm (2006) for more details.

classify jobs by type, such as sector, location, etc. For example, Dagsvik and Strøm (2006) apply a classification of job types according to sectors, namely the private and public sectors.

In a modeling context where "job" is allowed to be a decision variable, it is necessary to specify the choice set of available jobs in addition to the budget. The individual-specific sets of feasible jobs are endogenous in the sense that they are determined by market equilibrium conditions and/or by negotiations between unions and employers. However, to the individual agent, the set of job opportunities may be viewed as given. What complicates matters further is that these restrictions are typically latent, since the researcher usually has very little information about the agent-specific restrictions on job opportunities, or attributes of the chosen job. This introduces a new source of unobserved heterogeneity. The alternative approach we proposed here opens up for a more realistic treatment of this problem. It represents a powerful modeling strategy because it leads to an empirical framework that is flexible and practical to apply, and in provides, in our view, a better analogy to crucial features of the "true" choice setting. Unlike in standard models, it is easy to account for latent choice restrictions in our approach, and accordingly we can accommodate peaks in the hours of work distribution due to demand constraints. As will be explained below, it follows that our modeling framework can, formally, be viewed as an extension of the conventional discrete choice approach in the literature. It is common practice in some applied work to introduce dummy variables in the model specification to account for observed concentration of hours of work. This type of approach is purely *ad hoc* and hard to interpret in a structural context. In contrast, our alternative approach offers a theoretical rationale and interpretation for this practice. Our framework is also consistent with a story where the observed concentrations of hours are due to *both* preferences and constraints. Thus, if the purpose of the model is to simulate the effect of policy changes of the budget constraint this implies that one does not need to separate preferences from restrictions on the set of feasible working hours, provided these restrictions are kept fixed under policy changes.

A rather difficult and undercommunicated problem in behavioral sciences is the issue of functional form. Specifically, in most quantitative behavioral models the absence of theoretical justification for the choice of functional form and the distribution of unobservables is striking. The standard approach in this context is to select a flexible family of parametric or semi-parametric specifications, usually on grounds of convenience, and then proceed by using statistical inference methods to select a suitable specification within the a priori selected class of specifications. This issue is also closely related to the identification problem because identification often hinges on a selected family of functional forms. Unfortunately, it is in general insufficient to rely solely on statistical inference theory as a strategy for determining functional form and distributional properties of the error terms of behavioral models. The reason for this is that the class of possible model specifications is very large. Therefore, without theoretical principles almost any form is *a priori* possible and the correct one is difficult to determine because of data limitations, unobserved variables and measurement errors.

In a companion paper (Dagsvik and Jia, 2006) we report results of the empirical application and we discuss issues related to practical application of the modeling framework as regards estimation and simulation of policy reforms.

Although several novel aspects are being discussed extensively in this paper, our modeling approach is in several ways still rather simplistic. Perhaps the most important shortcoming is that it is a purely static approach which ignores possible endogenous savings decisions and how labor supply behavior depends on changes in pension rules, interest rates and wage profiles over the life cycle. For contributions on various aspects on intertemporal labor supply modeling see for example Blundell and MaCurdy (1999), Hyslop (1999), Keane and Wolpin (2001) and French (2005).

The paper is organized as follows. In Section 2, we present the basic modeling framework. In Section 3, we discuss further development of the basic framework. Specifically, this section discusses two different approaches for dealing with unobserved heterogeneity in the choice sets of jobs. In Section 4 we discuss functional form issues and Section 5 deals with the relationship to other approaches.

2. The basic modeling framework

In this section, we present the basic structure of our modeling approach. In contrast to the traditional approach in which the agent is restricted to have preferences solely over combinations of total consumption and hours of work we allow the agent to have preferences over total consumption, hours of work and nonpecuniary job attributes, such as the nature of the job-specific tasks to be performed, and location of the workplace, etc.

Let $U(C,h,z)$ be the (ordinal) utility function of the household, where *C* denotes household consumption (disposable income), and *h* is hours of work. The positive indices, $z = 1,2,...,$ refer to market opportunities (jobs) and $z = 0$ refers to the nonmarket alternative. For a market opportunity (job) *z*, associated hours of work and wage rate are assumed fixed and equal to $(H(z),W(z))$. In this section, we will assume that the hours of work and wage rate take only discrete values in a given set. A version with continuous hours and wage rates will be discussed in Section 3. Let *D* be the set of possible hours of work and *G* be the set of possible values of wage rate.

The utility function is assumed to have the structure

(2.1) $$U(C,h,z)=v(C,h)\varepsilon(z),$$

for $z=0,1,2,...,$ where $v(\cdot)$ is a positive deterministic function and $\{\varepsilon(z)\}$ are positive random taste shifters. The random taste shifters are assumed to account for unobservable individual characteristics and nonpecuniary job-type attributes that affect utility, and hence will vary both across households and job opportunities.

For given hours and wage rates, *h* and *w*, the economic budget constraint is represented by

(2.2) $$C=f(hw,I),$$

where *I* is nonlabor income, *C* is (real) disposable income and $f(\cdot)$ is the function that transforms gross income into after-tax household income. The function $f(\cdot)$ can in principle capture all details of the tax and benefit system.

For simplicity, we use the notation

(2.3) $$\psi(h,w,I)\equiv v\big(f(hw,I),h\big).$$

The term $\psi(h,w,I)$ is the representative utility of jobs with hours of work *h*, a given wage rate *w* and nonlabor income *I*.

In addition to (2.2), there are restrictions on the set of available market opportunities faced by a specific worker. This is because there are job types for which the worker is not qualified and there may be variations in the set of job opportunities for which he or she is qualified. Let $B(h,w)$ denote the agent's set of available jobs with hours of work and wage rate (h,w); that is, this set contains those jobs *z* for which $H(z)=h$ and $W(z)=w$. Let $m(h,w)$ be the number of jobs in $B(h,w)$. There is only one nonmarket alternative, so that $m(0,0)=1$. The choice sets $\{B(h,w)\}$ are unobserved to the researcher. Prior to job search, the individual-specific choice set of jobs may even be unknown to the agent and may be revealed through the search process in which the agent learns gradually about his or her (equilibrium) choice set. See Dagsvik (2000) for details of the interpretation of choice sets that are unknown to the agents prior to search. The random error terms $\{\varepsilon(z)\}$ are assumed to be independent and identically distributed (i.i.d.) across jobs and individuals with type I extreme value distribution[2] in the terminology of Resnick (1987).

[2] The cumulative distribution function is equal to exp(−1/x), defined for positive values of x. Note that the error terms being distributed according to the c.d.f. exp(−1/x) in a multiplicative formulation of the utility function is equivalent to the error terms being distributed according to the c.d.f. exp(−exp(−x)) in the corresponding and equivalent additive utility formulation. This follows by taking the logarithm of (2.1).

This particular distribution function is consistent with the property that the choice of jobs satisfies the assumption of independence from irrelevant alternatives (IIA), Luce (1959). Recall that the basic underlying intuition of the IIA assumption is that the agent's ranking of job opportunities from a subset, *B* (say), within the choice set of feasible jobs with given job-specific hours of work and wage rate, does not change if the choice set of feasible jobs is altered. For further interpretation and discussion of the IIA assumption we refer to Section 4.1 below.

Similarly to standard results in discrete choice theory (McFadden, 1984), it follows that the probability that a *specific* job, *z* (say), within *B*(*h,w*) is chosen, is given by

$$P\Big(\psi(h,w,I)\varepsilon(z)=\max_{x\in D,y\in G,}\max_{k\in B(x,y)}\big(\psi(x,y,I)\varepsilon(k)\big)\Big)=\frac{\psi(h,w,I)}{\sum_{x\in D,y\in G,}\sum_{k\in B(x,y)}\psi(x,y,I)}$$

$$=\frac{\psi(h,w,I)}{\psi(0,0,I)+\sum_{x\in D,x>0}\sum_{y\in G}\psi(x,y,I)m(x,y)}.$$

Let $\varphi(h,w\,|\,I)$ denote the probability that the agent chooses a particular job with offered hours *h*, wage rate *w*, given nonlabor income *I* (and individual characteristics). This probability is equal to the probability of choosing *any* job within *B*(*h,w*), and is thus obtained by summing the choice probabilities above over all jobs in *B*(*h,w*), yielding

(2.4)
$$\varphi(h,w\,|\,I)=\sum_{z\in B(h,w)}P\Big(\psi(h,w,I)\varepsilon(z)=\max_{x\in D,y\in G,}\max_{k\in B(x,y)}\big(\psi(x,w,I)\varepsilon(k)\big)\Big)$$

$$=\sum_{z\in B(h,w)}\frac{\psi(h,w,I)}{\sum_{y\in G,}\sum_{x\in D,x>0,}\sum_{z\in B(x,y)}\psi(x,y,I)+\psi(0,0,I)}=\frac{\psi(h,w,I)m(h,w)}{\psi(0,0,I)+\sum_{y\in G,}\sum_{x>0,x\in D}\psi(x,y,I)m(x,y)}$$

for $h,w>0,$ and

(2.5)
$$\varphi(0,0\,|\,I)=\frac{\psi(0,0,I)}{\psi(0,0,I)+\sum_{y\in G,}\sum_{x>0,x\in D}\psi(x,y,I)m(x,y)}$$

for $h=0.$ The resulting expression is a choice model that is analogous to a multinomial logit model with representative utility terms $\{\psi(h,w)\}$, weighted by the frequencies of available jobs, $\{m(h,w)\}$. Note that it is a consequence of our distributional assumptions of the stochastic error term in the utility function that the respective numbers of available latent jobs, $\{m(h,w)\}$, represent a set of *sufficient* statistics for the corresponding choice sets. Unfortunately, the $\{m(h,w)\}$ are not directly observable, but under specific assumptions, one can identify $m(h,w)$ and $\psi(h,w,I)$ and estimate their parameters. For the sake of interpretation, and with no loss of generality, write $m(h,w)=\tilde{\theta}\tilde{g}(h,w)$, where

$$\tilde{\theta}=\sum_{y\in G,}\sum_{x\in D, x>0} m(x,y),$$

and $\tilde{g}(h,w)=m(h,w)/\tilde{\theta}$. The interpretation of $\tilde{g}(h,w)$ is as the fraction of available jobs (available to the agent) with offered hours of work and wage rates equal to (*h, w*), whereas the parameter $\tilde{\theta}$ is the total number of jobs available to the agent. In this sense this setup is analogous to the formulation in Tummers and Woittiez (1990), Dickens and Lundberg (1993) and Bloemen (2000), who specify joint offer distributions of hours and wage rates to represent the agent's choice sets. However, so far the latent choice sets of jobs cannot be interpreted as random in the sense that they vary across agents. This issue will be developed below.

3. Further development of the basic framework

The framework presented above is a version of a random utility model modified to account for aggregation of latent alternatives. Early versions of this framework in the context of labor supply analysis were applied by Dagsvik and Strøm (1988) and Aaberge, Dagsvik and Strøm (1995). Analogous versions in other modeling contexts have been discussed by Ben-Akiva and Watanatada (1981) and Ben-Akiva et al. (1985). It is however desirable to extend the basic framework above in order to make the approach more realistic in the context of empirical applications and simulation of policy reforms. In the following we discuss two different approaches for dealing with unobserved heterogeneity in the choice sets of jobs. The first one is traditional in the sense that the parameters representing the choice sets are viewed as random effects. The second one is based on a Poisson process representation, similarly to Dagsvik (1994) and Dagsvik and Strøm (2006).

3.1. Unobserved heterogeneity in choice sets and in inter - and intra individual-specific tastes

In the preceding analysis, we treated the terms that represent the size of the choice sets as a constant across observationally identical households. Thus, potential unobserved heterogeneity in choice sets across households is ignored. In fhe following we present an approach which accounts for this unobserved heterogeneity by representing variations in choice sets and in inter - and intra individual-specific tastes as random variables.

The traditional interpretation of the taste-shifters $\{\varepsilon(\mathrm{z})\}$ is that they solely account for nonsystematic variation in preferences across agents and jobs due to unobservables, but are viewed as perfectly known to the respective agents. Another interpretation views tastes as random also to the agent himself in the sense that when presented with repetitions of identical choice settings, the agent may make different choices on each occasion, cf. Morrison (1962), Quandt (1956), and Tversky (1969). The reason for this is that agents may have insufficient information and experience (bounded rationality) with the choice alternatives and accordingly find it difficult to make a precise assessment of their utility onces and for all. Of course, from the observing researcher's perspective the formulation (2.1) above is consistent with both interpretations. In the development that follows we shall make explicit use of the two different interpretations of the sources of randomness in tastes.

To this end, assume that the random error terms introduced in (2.1) are extended to

$$\tilde{\varepsilon}(\mathrm{z}) = \eta(\mathrm{z})\kappa(\mathrm{z}) \tag{3.1}$$

where the terms $\{\kappa(\mathrm{z})\}$ are random variables that represent the values of the unobservable aspects of the alternatives that are perfectly known to the agent. For a given job z, $\kappa(\mathrm{z})$ is constant for each agent across identical "choice experiments" but varies across agents. Without loss of generality, we normalize $\{\kappa(\mathrm{z})\}$ such that $\kappa(0) = 1$. In other words, $\kappa(\mathrm{z})$ represents the value of non-pecuniary aspects of job z relative to the non-market opportunities. The terms $\{\eta(z)\}$ are random variables that represent the values of the "uncertain" aspects of the alternatives. By uncertain aspects we mean in this context aspects that the agent finds difficult to evaluate, either because of the agent's psychological state of mind that may vary from one moment to the next or because he or she has insufficient information about features of the jobs. For a given z, $\eta(\mathrm{z})$ may therefore vary across identical choice settings for each agent. This type of intra-individual randomness has a long tradition in psychology, dating back to Thurstone (1927). Consistent with IIA, assume that $\{\eta(z)\}$ are i.i.d. across jobs with type-I extreme value distribution, as in Section 2. We also assume that $\{\kappa(z)\}$ and $\{\eta(\mathrm{z})\}$ are independent. Let

$$\mu(h,w) = \sum_{z \in B(h,w)} \kappa(z).$$

Note that although from the agent's point of view, the function $\mu(h,w)$ is known, it is perceived as random by the observing econometrician. However, it has no longer the interpretation as the number of feasible job opportunities. Note also that the assumption in (3.1) implies that $\mu(h,w)$ is unbounded, in contrast to $m(h,w)$, defined in Section 2, and that the number of $\kappa(z)$ that appear in the sum above may be random in the sense that it may vary over agents because $B(h,w)$ may vary across agents. Furthermore, let $\varphi\big(h,w \mid I,\{\mu(x,y), x\in D, y\in G\}\big)$ denote the conditional probability of supplying h hours of work given the wage rate, nonlabor income and given the terms $\{\mu(h,w), h\in D, w\in G\}$. Similarly to (2.4), it follows immediately that the conditional density of supplied hours of work, given $\{\kappa(z)\}$, has the structure

$$\sum_{z\in B(h,w)} P\Big(\psi(h,w,I)\varepsilon(z)=\max_{y\in G,x\in D,x>0,}\max_{k\in B(x,y)}\big(\psi(x,y,I)\varepsilon(z)\big)\mid\{\kappa(z)\}\Big)$$

$$(3.2)\qquad =\varphi\big(h,w\mid I,\{\mu(x,y),x\in D,y\in G\}\big)=\frac{\psi(h,w,I)\mu(h,w)}{\psi(0,0,I)+\sum_{y\in G,}\sum_{x\in D,x>0}\psi(x,y,I)\mu(x,y)}.$$

Note that the properties of the choice set and the aspects that are known to the agent are fully represented by $\{\mu(h,w)\}$ in the model. In other words, the set $\{\mu(h,w), h\in D, w\in G\}$ represents a sufficient set of random variables for the latent choice sets $\{B(h,w), h\in D, w\in G\}$. Evidently, when $\kappa(z)=1$, (3.2) reduces to (2.4), although $\mu(h,w)$ is now random instead of fixed. From (3.2) it follows that the unconditional choice probability of working h hours is given by

$$(3.3)\qquad \varphi\big(h,w\mid I\big)=E\varphi\big(h,w\mid I,\{\mu(x,y),x\in D,y\in G\}\big),$$

where the last expectation is taken with respect to $\{\mu(h,w), h\in D, w\in G\}$. Assume furthermore that $\mu(h,w)=\tilde{\theta}\tilde{g}(h,w)\omega(h,w),$ for positive h and w, and $\mu(0,0)=\omega(0,0),$ where $\{\tilde{\theta}\tilde{g}(h,w)\}$ have the same interpretation as above and $\{\omega(h,w)\}$ are i.i.d. random terms that are independent of the deterministic parts of the utility function.

A challenging issue is how to characterize the distribution of the terms $\{\omega(h,w), h\in D, w\in G\}$. Our approach to this end is to postulate plausible properties we believe this distribution should possess and subsequently derive the implications. The properties we postulate are the following: (i) $\omega(h,w)>0$; (ii) for any hours of work, h_1 and h_2, and nonnegative constants, b_1 and b_2, $b_1\omega(h_1,w_1)+b_2\omega(h_2,w_2)$ has the same distribution as $\tau\omega(h_1,w_1),$ where τ is a positive constant that may depend on h_1, w_1 h_2, w_2, b_1 and b_2; (iii) the random variables $\omega\big(h_1,w_1\big)$ and $\omega\big(h_2,w_2\big)$ are i.i.d. The motivation for (i) is obvious: unless this condition is satisfied, for some hours of work, the conditional choice probabilities would be zero or negative. Conditions (ii) and (iii) mean that for any positive $h_1,w_1,h_2,w_2,...,h_r,w_r,$ the distribution of the conditional aggregate choice probabilities

$$\sum_{k=1}^{r} \varphi\left(h_k, w_k \mid I, \{\kappa(z)\}\right),$$

(which are random variables because they depend on $\{\kappa(z)\}$ through $\{\mu(h,w), h \in D, w > 0\}$), across unobservable choice sets, belongs to the same family of distributions as the conditional choice probabilities, $\varphi\left(h_k, w_k \mid I, \{\kappa(z), z = 0,1,..\}\right)$). In other words, requirement (ii) implies that the distribution of the conditional choice probabilities is *invariant* under aggregation of alternatives (combination of hours of work and wage rates). The motivation for propterty (ii) is that since the aggregation level within the total set of feasible hours and wage rates is somewhat arbitrary, it seems intuitive that the distributional properties of the model should not depend critically on the partition of the set of feasible hours and wage rates into aggregate alternatives. See also Section 4.2 for more discussion on invariance assumptions.

It can be demonstrated that the postulated assumptions imply that the distribution of $\omega(h,w)$ is strictly Stable.[3] Moreover, when $\{\omega(h,w)\}$ are independent and distributed according to a strictly Stable distribution, it is shown in Appendix A that

$$(3.4) \qquad \varphi(h,w \mid I) = E\left(\frac{\psi(h,w,I)\mu(h,w)}{\sum_{y \in G,\, x \in D} \sum \psi(x,y,I)\mu(x,y)} \right) = \frac{\psi(h,w,I)^{\alpha}\theta g(h,w)}{\psi(0,0,I)^{\alpha} + \sum_{y \in G,\, x \in D, x>0} \sum \psi(x,y,I)^{\alpha}\theta g(x,y)},$$

for $h > 0,$ and similarly for $h = 0,$ where the expectation is taken with respect to $\{\omega(h,w)\},$ $\alpha \le 1,$ is a positive parameter, $g(h,w) = \tilde{\theta}^{\alpha}\tilde{g}(h,w)^{\alpha}/\theta,$ and

$$\theta = \tilde{\theta}^{\alpha} \sum_{x \in D,\, y \in G} \sum \tilde{g}(x,y)^{\alpha}.$$

Since the jobs in our context are unobservable, the measures $\tilde{\theta}\tilde{g}(h,w)$ and $\theta g(h,w)$ are equivalent in that they both provide sensible interpretations of the consentration of latent market opportunities. We shall call $\theta g(h,w)$ the *opportunity measure* and $g(h,w)$ the *opportunity density*. The opportunity density is equivalent to a probability mass function, namely the probability that, the set of available jobs to a randomly selected agent, contains a job with hours of work and wage rate (*h, w*). This means that the opportunity density has the interpretation as the offered distribution of hours and wages, similarly to the formulation of Tummers and Woittiez (1990), Dickens and Lundberg (1993) and

[3] Recall that the class of Stable distributions represents a generalization of the normal distributions. In particular, a general version of the central limit theorem yields the class of Stable distributions; see for example, Embrechts, Klüppelberg and Mikosch (1997) for a description of Stable distributions.

Bloemen (2000). The interpretation of θ can be extended to include fixed cost. To realize this, assume that a positive parameter *c*, representing the utility (disutility) of fixed cost, enters multiplicatively in the utility function given in (2.1) for positive hours of work. Then, evidently, the structure of the choice probabilities above remains the same apart from θ which now transforms to θc^{α}.

Thus, we have seen that our particular approach to random choice sets imply that the structure of the choice probabilities is invariant under aggregation across unobserved choice sets (with suitable reinterpretation of the opportunity measure), except for a power transformation of the systematic part of the utilities. In other words, we have demonstrated that the structure of the labor supply choice probabilities given in Section 2 is consistent with the stochastic choice sets of feasible jobs provided that the systematic part of the utility function *v* has a functional form that is invariant under increasing power transformations. This aggregation result comes at the cost of the assumption of $\omega(h,w)$ being independent of $\omega(h',w')$ when $(h,w)\neq(h',w')$. However, one can in fact demonstrate that similar invariance aggregation results holds also when $\omega(h,w)$ and $\omega(h',w')$ are not independent (Dagsvik, 2008).

Let us for a moment relate this particular random effect approach to the traditional approach to random effects. This consists in specifying an arbitrary distribution for $\{\omega(h,w)\}$, for example a lognormal distribution. This would imply a complicated expression for the choice probabilities since no closed form solution exists for the expectation in (3.4) in this case. Moreover, the choice of the lognormal distribution is, from a theoretical perspective, totally arbitrary in this context. In contrast, the approach outlined above has the advantage of both having theoretical support and resulting in very convenient expression for the choice probabilities. We shall later discuss the inclusion of additional random effects that are *not* alternative specific.

3.2. A general framework with random choice sets generated by a Poisson process

In contrast to the random effect approach discussed in the preceeding section, Dagsvik (1994) proposed another framework for dealing with stochastic choice sets that accommodates unobserved heterogeneity in choice sets. This framework is based on a particular nonhomogeneous multidimensional Poisson process representation. Dagsvik and Strøm (2006) have applied this framework in the context of labor supply modeling. In this subsection we shall present an extension of this approach.

Assume now that the taste-shifters $\{\varepsilon(z), z=1,2,...\}$, of the available jobs are points of a Poisson process on $(0,\infty)$ with intensity $\lambda(\varepsilon)$, where $\lambda(\varepsilon)$ is a suitable positive and decreasing function. Thus, the probability that a job with taste-shifter within $(\varepsilon,\varepsilon+d\varepsilon)$ is equal to $\lambda(\varepsilon)d\varepsilon$.

Moreover, the probability that there are more that one job with taste-shifters within $(\varepsilon, \varepsilon + d\varepsilon)$ is negligible. The expected number of jobs available with $\varepsilon > 1/x$, for some nonnegative *x*, is equal to

(3.5) $$\Lambda(x) = \int_{1/x}^{\infty} \lambda(\varepsilon) d\varepsilon.$$

The integral above is assumed to be finite for finite *x*, but may tend towards infinity when *x* tends towards infinity. The interpretation of $\Lambda(x)$ is as the total expected number of jobs available for which $\varepsilon(z) > 1/x.$

Assume the corresponding hours of work and wage rates $\{(H(z), W(z))\}$ associated with the jobs are independent draws from the joint density *g*(*h,w*). Then, it follows from Resnick (1987, p.135) that the points $\{(H(z), W(z), \varepsilon(z)), z = 1, 2, ...,\}$ are realizations of a non-homogeneous three-dimensional Poisson process on $D \setminus \{0\} \times G \times (0, \infty)$ with intensity measure $g(h, w)\lambda(\varepsilon),$ where the sets *D* and *G* now are allowed to be continuous. This intensity provides a complete representation of the Poisson process in the sense that it governs the corresponding probability distribution of the Poisson points in Ω. See for example Resnick (1987) for definitions and properties of Poisson processes. Recall that similarly to the one-dimensional case, if the Poisson process is homogeneous the points are randomly but evenly dispersed on Ω. In contrast, the non-homogeneous Poisson process allows for uneven distribution of points in the sense that there are, on average, higher concentration of points in some parts of Ω than in other parts. This allows us to account for the possibility that there may be fewer jobs available with high values of the taste shifters $\{\varepsilon(z)\}$ than jobs with low values of $\{\varepsilon(z)\}$. Moreover, it is assumed that $\{\varepsilon(z), z = 1, 2, ...\}$ are distributed independently of $\varepsilon(0)$. Finally, we assume that $\varepsilon(0)$ has c.d.f. $\exp(-1/x),$ for positive x. It follows that the probability that there is a job with $H(z) \in (h, h + dh), W(z) \in (w, w + dw), \varepsilon(z) \in (\varepsilon, \varepsilon + d\varepsilon)$ is available to the individual equals $g(h, w) dh dw \lambda(\varepsilon) d\varepsilon.$ Moreover, the multiplicative form of the intensity measure means that $\{(H(z), W(z)), z = 1, 2, ...\}$ are independent of $\{\varepsilon(z), z = 0, 1, 2, ...\}$. A more general version of the intensity allows the taste-shifters to depend on the offered hours and wages. Specifically, this will be the case if λ is allowed to depend on (*h,w*), i. e., $\lambda(\varepsilon)$ is replaced by $\lambda(\varepsilon \mid h, w),$ and consequently $\Lambda(x)$ is replaced by $\Lambda(x \mid h, w).$ The last specification may be of interest in applications where one believes there may be correlation between job-specific nonpecuniary attributes, offered wage rates and hours of work. We are now ready to express the probability distribution of realized hours and wages, including the probability of not working. Let $\Phi(h,w\,|\,I)$ be the joint cumulative distribution of realized hours and wages that follow from utility maximizing behavior, i.e.,

$$(3.6)\qquad \Phi(h,w\,|\,I) \equiv P\left(\max_{z \atop H(z)\le h, W(z)\le w}\left(\psi\left(H(z),W(z),I\right)\varepsilon(z)\right) = \max_z\left(\psi\left(H(z),W(z),I\right)\varepsilon(z)\right)\right),$$

and let $\varphi(h,w\,|\,I)$ be the corresponding density. In Appendix B it is proved that the conditional probability density $\tilde{\varphi}(h,w\,|\,U,I)$ of the chosen hours of work and wage rate, given *I* and the utility level *U*, is given by

$$(3.7)\qquad \tilde{\varphi}(h,w\,|\,U,I) = \frac{\Lambda'\left(\frac{\psi(h,w,I)}{U}\,|\,h,w\right)\psi(h,w,I)g(h,w)}{\frac{\psi(0,0)}{U} + \int\limits_{x\in D,x>0}\int\limits_{y\in G}\Lambda'\left(\frac{\psi(x,y,I)}{U}\,|\,h,w\right)\psi(x,y,I)g(x,y)dxdy},$$

for $h>0, w>0$. Furthermore, the (indirect) utility *U* has c.d.f. given by

$$(3.8)\qquad P(U\le u\,|\,I) = \exp\left(-\frac{\psi(0,0,I)}{u} - c\int\limits_{x\in D,x>0,}\int\limits_{y\in G}\Lambda\left(\frac{\psi(x,y,I)}{u}\,|\,h,w\right)g(x,y)dxdy\right).$$

Hence, it follows that the (uncompensated) choice probability density is given by

$$(3.9)\qquad \varphi(h,w\,|\,I) = E_U\tilde{\varphi}(h,w\,|\,U,I),$$

where E_U is the expectation with respect to *U*. Thus, with the assumptions given above, the distribution of realized hours of work becomes a continouous one, in contrast to the setup in the previous sections.

In the special case with $\Lambda(x)=\theta x,$ we obtain the model obtained in Dagsvik and Strøm (2006), given by

$$(3.10)\qquad \varphi(h,w\,|\,I) = \frac{\theta\psi(h,w,I)g(h,w)}{\psi(0,0,I) + \theta\int\limits_{x\in D,x>0,}\int\limits_{y\in G}\psi(x,y,I)g(x,y)dxdy}$$

for $h>0, w>0,$ and similarly for $h=w=0$. Recall that θx is the expected number of available jobs with taste-shifters greater than 1/*x*. In particular, θ is the expected number of available jobs with taste-shifters greater than one, and therefore one can interpret θ as a measure of the total amount of job opportunities available.

If instead, hours of work and wage rates are drawn from a discrete distribution one gets a discrete distribution of realized hours and wage rates. In this case (h,w) belongs to a finite set or countable set, and the structure of the choice probabilities have the same structure as above with the integrals replaced by sums.

Note that whereas (3.10) satisfies the Independent from Irrelevant Alternatives assumption (IIA), this is not necessarily so in the more general case given in (3.6) to (3.8). It follows from Dagsvik (1994) that unless $\Lambda(x)$ is a power function the resulting choice probabilities will *not* satisfy IIA. The modeling framework expressed above in (3.6) to (3.8) thus generalizes the model considered in Dagsvik and Strøm (2006) and also the approach proposed by Bloemen (2000).

For the sake of interpretation and relation to the literature on random utility models, consider for a moment the choice of hours (that is job-specific hours of the chosen job) where we for simplicity assume that the wage rate is purely individual-specific. Let $\tilde{U}(h)$ be the (indirect) utility of the preferred job among jobs with h hours of work. It is easily verified that one can write $\tilde{U}(h) = \psi(h,w)\tilde{\varepsilon}(h),$ where $\tilde{\varepsilon}(h)$ is a positive random error term which is independent of $\psi(h,w),$ and with c.d.f.

$$(3.11) \qquad F_h(y) \equiv P\left(\tilde{\varepsilon}(h) \le y\right) = \exp\left(-\Lambda\left(\frac{1}{y} \mid h\right) g(h)\right),$$

for positive y. This means that one can interpret the framework above as a conventional independent random utility model, where the relationship between the c.d.f. of $\tilde{\varepsilon}(h),$ $g(h)$ and $\Lambda\left(x \mid h\right)$ is given in (3.11). In the case where $\Lambda\left(x \mid h\right)$ depends on h the distribution function in (3.11) will be alternative specific even if the density g is uniform. It is known that the independent random utility model with *identically* distributed utilities will satisfy IIA approximately, but this is not necessarily the case when the distributions are alternative specific. For the general case Sattath and Tversky (1976) have shown that the independent random utility model (with absolutely continuous c.d.f.) imply a (non-parametrically) testable property called the multiplicative inequality.

In this section we have seen, similarly to the previous section, how one can obtain convenient expression for choice probabilities for continuous choices that are consistent with stochastic choice sets. The treatment here differs from the conventional one in that we did obtain the formulas above by applying a random effect type of formulation.

4. Functional form issues

The modeling framework presented above offeres a very general and flexible approach to labor supply modeling. However, it is so far of limited interest unless additional assumptions are imposed as to the structure of the utility function and the opportunity measure. In our model, the observed accepted wage and hours of work is a result of both the preference (utility function) and job offer distribution (the opportunity measure). Without further assumptions the modeling framework given in (2.4) and (2.5) is not identified[4]. In this context the issue of functional form is particularly important. A delicate and often neglected issue in most quantitative behavioral models is the lack of theoretical justification for the choice of functional form and the distribution of unobservables. The standard approach in this context is to select a flexible family of parametric or semi-parametric specifications, usually on grounds of convenience, and then proceed by using statistical inference methods to select a suitable specification within the a priori selected class of specifications. This issue is also closely related to the identification problem because identification often hinges on a selected family of functional forms. Unfortunately, it is in general insufficient to rely solely on statistical inference theory as a strategy for determining functional form and distributional properties of the error terms of behavioral models. The reason for this is that the class of possible model specifications is very large. Therefore, without theoretical principles almost any form is *a priori* possible and the correct one is difficult to determine because of data limitations, unobserved variables and measurement errors. A full nonparametric approach is not possible in practice, because it requires that the researcher has access to an unlimited set of data covering behavioral responses that correspond to every possible counterfactual and relevant policy regime. For the same reason the value of statistical testing of parametric functional form assumptions is limited because one usually has at most access to data on just a few alternative policy regimes. Thus, unless one is able to justify the choice of functional form of the behavioral relations policy implications may be misleading. As Simon (1986)[5] has pointed out, many conclusions that have been drawn in the literature about the way in which the economy operates depend crucially on ad hoc assumptions about the functional form of the agents' utility functions.

In the labor supply literature several researchers have applied flexible functional forms such as translog or polynomial specifications of the utility function. In addition to the fact that these families of functional forms are ad hoc, these specifications are problematic because one cannot guaranty

[4] Recently, several authors have suggested using subjective information such as desired hours of work to estimate labor supply models. See for example, Bloemen (2008) and references therein. In principal, if one can identify the preferences using desired hours of work first, one can then identify the job offer distributions based on actual observed hours of work.

[5] Simon (1986):"Contemporary neoclassical economics provides no theoretical basis for specifying the shape and content of the utility function, and this gap is very inadequately filled by empirical research using econometric techniques. The gap is important because many conclusions that have been drawn in the literature about the way in which the economy operates depend on assumptions about consumers' utility function."

global quasi-concavity and monotonicity (see Dagsvik and Strøm, 2006). As we show in our companion paper, the estimated specification based on a quadratic polynomial representation of the deterministic part of the utility function yields decreasing utility of leisure in some intervals, which is unacceptable from a theoretical point of view.

The building blocks in our framework consist of the distribution of the random terms of the utility function, the deterministic part of the utility function and the opportunity density. We now turn to a discussion and justification of the first two components, namely the distribution of the random- and the deterministic terms of the utility function.

4.1. The distribution of the random part of the utility function and Independence from Irrelevant Alternatives (IIA)

In our approach we have used IIA to motivate the choice of extreme value distributed taste-shifters. Some researchers seem, a priori, to consider the IIA assumption as severely restrictive and therefore unacceptable in an empirical modeling context. Although IIA may be restrictive in some settings, it is our opinion that it nevertheless represents a powerful principle of *probabilistic rationality*, and therefore is a natural point of departure for formulation stochastic models of choice behavior. As Luce (1977), pp. 232-233 puts it:

> "Perhaps the greatest strength of the choice axiom (IIA), and one reason it continues to be used, is as a canon of probabilistic rationality. It is a natural probabilistic formulation of K.J. Arrow's famed principle of the independence of irrelevant alternatives, and as such it is a possible underpinning for rational, probabilistic theories of social behavior. Thus, in the development of economic theory based on the assumption of probabilistic individual choice behavior, it can play a role analogous to the algebraic rationality postulates of the traditional theory."

We shall now show that IIA postulated in our approach can in fact also be related and compared to the traditional approach based on conventional micro theory, and which is discussed in Section 5.1. Specifically, we shall show that in some sense the traditional algebraic approach, extended to incorporate population heterogeneity, is in fact more restrictive than the probabilistic approach based on IIA. To realize this, consider the setting described in Section 5.1, extended to the case with quantity restrictions on hours of work. To compare the two approaches it is necessary to introduce quantity constraints in the algebraic approach that correspond to the choice set in the probabilistic approach. Assume as above that the set of possible hours is a continuum *K* and let *A* denote a subset of *K.* Let $\tilde{h}_i(A) = F_A\left(w_i, I_i, \varepsilon_i\right)$ denote the corresponding constrained labor supply function, constrained to the choice set of hours A, where ε_i is an individual specific random variable representing tastes and F_A is a function that follows from the underlying utility specification. Let *B* be a subset of *A*. Define

(4.1) $$P_A(B) = P\left(\tilde{h}_i(A) \in B\right).$$

The empirical counterpart to $P_A(B)$ is the fraction of workers that choose hours of work within *B*, subject to the quantity constraints represented by the choice set *A*. It follows from quasi-concavity and monotonicity of the utility function that for two overlapping choice sets, A_1 and A_2, that

(4.2) $$P_{A_1}(B) = P_{A_2}(B)$$

provided *B* belongs to the interior of $A_1 \cap A_2$. This is due to the fact that only the point of tangency between the budget line and the indifference curve matters for the determination of the supply of hours. In other words, when *B* belongs to the interior of the choice sets the theory predicts that the choice is independent of *A*. Thus, the conventional theory yields restrictions that are similar to IIA and appear even more restrictive than IIA, since IIA only predicts that $P_A(B_1)/P_A(B_2)$ is independent of *A*, where B_1 and B_2 are sets belonging to *A*. An equivalent statement of IIA is that

(4.3) $$\frac{P_{A_1}(B_1)}{P_{A_2}(B_1)} = \frac{P_{A_1}(B_2)}{P_{A_2}(B_2)},$$

for $A_1, A_2 \subset B_1 \cap B_2$, which of course is a weaker condition than $P_{A_1}(B)/P_{A_2}(B) = 1,$ which is equivalent to (4.2). Thus, when critizising IIA in this context one should be aware of the fact that the algebraic approach implies (in the particular sense defined above) a *stronger* restriction than IIA, as demonstrated above.

Nevertheless, one cannot rule out a priori that IIA may be too restrictive in our context. Fortunately, there are several practical approaches to relax the IIA. One is the so-called Mixed Multinomial Logit type approach, see McFadden and Train (2000). Another is the Nested Multinomial Logit type approach, (or more generally, the Generalizes Extreme Value model, see McFadden, 1984). Below we show how a particular Nested Multinomial Logit type of model can be formulated. Still, we believe that it is of importance, beyond the grounds of convenience, to maintain IIA as the basic underlying theoretical assumption against which extended versions may be tested precisely because of the interpretation of IIA as probabilistic rationality, summarized in the citation above by Luce (1977).

In the final part of this section we shall demonstrate that our model can readily be extended to Nested Multinomial Logit type models. Suppose now that the correlation between the random error terms in the utility function for positive hours of work alternatives is allowed to be positive. A rationale for this is that the agent may have taste for work, or equivalently, that tasks to be performed for jobs in the choice set may be similar and influence the taste for work. In the Appendix C we show that under typical assumptions about the correlation pattern within a Generalized Extreme Value framework (McFadden, 1984) the conditional choice probability in (3.4) extends to

$$\varphi(h,w\,|\,I)=\frac{\psi(h,w,I)^{1/\rho}g(h,w)\theta^{\rho}\left(\sum_{x\in D,y\in G,x,y>0,}\psi(x,y,I)^{1/\rho}g(x,y)\right)^{\rho-1}}{\psi(0,0,I)+\theta^{\rho}\left(\sum_{x\in D,y\in G,x,y>0,}\psi(x,y,I)^{1/\rho}g(x,y)\right)^{\rho}}, \tag{4.4}$$

Where $\rho\in(0,1]$ is a correlation parameter. Specifically, $1-\rho^2$ can be interpreted as the correlation between the error term in the utility function between different positive hours of work alternatives. We note that in the special case with $\rho=1$, the expression in (4.4) reduces to the formula in (3.4).

4.2. Functional form of the deterministic part of the utility function

Within psychology and psychophysics theories have been developed with the purpose of justifying functional form on the basis of invariance principles, cf. Falmagne (1985) and Narens (2002). These principles are similar to certain invariance principles applied in physics. Many models in physics are typically invariant under uniform translation and rotation of the coordinate system. To this end, Dagsvik and Strøm (2006) have applied a typical approach in this tradition. Specifically, they postulate particular invariance properties to obtain a characterization of the functional form of the deterministic part of the utility function. We shall now briefly summarize their approach. For detailed and precise statements we refer to Dagsvik and Strøm (2006). See also Dagsvik, Strøm and Jia (2006), where an analogous approach is pursued. First, assume that the utility function is given as in (2.1) with random error terms $\{\varepsilon(z)\}$ that are i.i.d., and independent of the structural term $v(C,h)$. However, apart from the i.i.d. assumption no additional distributional assumptions are needed.

In this context it turns out to be convenient to consider the case in which the sets of feasible jobs are equal for each level of hours of work and wage rates. This represents no loss of generality, since preferences are assumed to be independent of the choice sets. The first invariance assumption states that if the fraction of workers that prefer modified diposable income and leisure combinations, (C_1,L_1) to (C_2,L_2) is less than the fraction of workers that prefer disposable income and leisure combinations, (C_1^*,L_1) to (C_2^*,L_2), then the same is true when the respective modified levels of disposable incomes are scale transformations of the original levels. The second invariance assumption is symmetric to the first one and states that a similar property holds when the role of modified disposable income and leisure are interchanged in the sense that suitable (modified) leisure levels are rescaled whereas the original modified disposable income levels are kept fixed. These invariance assumptions above capture the notion that when the individual basic needs (subsistence) are fulfilled then the absolute levels of quantities tend not to be essential, rather the individuals relate to relative consumption levels. Dagsvik and Strøm (2006) provide further discussion on the limitation of these invariance assumptions. The notion that relative stimuli levels matter (beyond some lower or upper level treshold) rather than absolute ones is supported by numerous stated preference experiments, see for example Stevens (1975). Dagsvik and Strøm (2004) demonstrate that under general regularity conditions, the above invariance assumptions imply that the systematic part of the household utility function has the form

(4.5) $$\log v(C,h) = \beta_1 \frac{\left(C^{\alpha_1} - 1\right)}{\alpha_1} + \beta_2 \frac{\left(L^{\alpha_2} - 1\right)}{\alpha_2} + \beta_3 \frac{\left(C^{\alpha_1} - 1\right)}{\alpha_1} \frac{\left(L^{\alpha_2} - 1\right)}{\alpha_2},$$

where $L = 1 - h/T$, and T is total time available after sleep and rest have been deducted, and with the usual convention that the Box-Cox transformation $(x^\alpha - 1)/\alpha = \ln(x)$ when $\alpha = 0$. To ensure that the function $v(C,h)$ is increasing in C and strictly decreasing in h and concave one must have that $\alpha_1 < 1, \alpha_2 < 1, \beta_1 > 0, \beta_2 > 0,$ and in addition β_3 is positive, or if negative, sufficiently small numerically. In the estimation procedure, these restrictions can be easily imposed *a priori* or be checked after the estimation.

Although the assumptions implying (4.5) have considerable intuitive appeal, it would certainly be desirable if further evidence in support of the invariance assumptions above could be provided. A great advantage with the approach outlined in this section is that the postulated invariance assumptions can be tested directly, and independent of the functional form implication in (4.5), by means of suitable Stated Preference (SP) survey data. Falmagne and Iverson (1985), and Dagsvik and Røine (2008) have developed appropriate statistical testing procedures to this end. These tests are *nonparametric* and formulated as inequalities within a binomial or multinomial setting. Recall that SP data allows the researcher to collect several observations for each individual under alternative conditions. Also, one can specify conditions similarly to controlled laboratory experiments, such that the maintained condition of choice set of jobs being constant for each individual under alternative economic budget constraints. Subsequently, one can test if the model specification selected on the basis of theoretical assumptions and supported by SP data is consistent with the available "real" market data. The advantage with this approach is that one avoids the controversial initial ad hoc step of selecting a family of a priori functional forms within which conventional statistical testing is carried out[6]. As mentioned above, one can instead test the invariance assumptions proposed above without specific a priori, and unjustified assumptions about functional form.

The approach discussed in this section is in our view promising but still only provides a partial solution of the functional form issue for the type of labor supply models discussed in this paper. However, a similar characterization and justification of the functional form of the opportunity density function remains a challenge for future research.

4.3. Specification of the opportunity measure

In this section, we discuss the structure of the opportunity measure. In the introduction we mentioned briefly that the latent sets of available jobs to the workers are endogenous. That is, although we have assumed that the agent's taste-shifters are (stochastically) independent of offered hours and wage rates, the distribution of wage rates and the opportunity density will depend on the *distribution* of the preferences due to equilibrium conditions. In other words, the market forces that regulate the balance between supply and demand, be it a market-clearing regime or not, are assumed to operate solely at the aggregate level. Consequently, the opportunity density depends on the production technologies of firms as well as on the contracts and wage-setting policies of unions and firms. It is beyond the scope of this paper to discuss fully how the distribution of opportunity measure, $\theta g(h,w)$ through market equilibrium adjustments, depends on the systematic part of the utility function, $\psi(\cdot)$. Dagsvik (2000) considers equilibrium conditions in a setting in which the labor market is viewed as a matching game in which workers and firms search in order to obtain the best possible match with a potential partner. However, it is too demanding to implement these conditions in our setting. We therefore use a reduced form specification of the opportunity measure. Note that under this reduced form specification, the estimated model can only be applied to simulate behavior conditional on the opportunity density.

As we discussed earlier, θ can be interpreted as the number of jobs that are feasible to the individual weighted by the utility of working, including the disutility of fixed cost of working. We assume that θ may depend on variables that represent the effect of schooling, experience and charaterize local labor market conditions, and possible variables that represent fixed cost of working, as discussed above.

---

[6] It is true that many economists remain sceptical towards stated preference data. This is because they typically believe that agents reveal their true preferences only under market like conditions in which financial incentives matter. However, in recent years researchers have studied the issue of validity of SP data and have concluded that in many circumstances SP data seem to be reliable. In any case, it is clear that one almost never obtains market data that are nearly so varied and detailed as data obtained from SP surveys. Clearly, unless one thinks that such data are worse than no data at all, the researcher can fruitfully use these data to establish support for particular functional form properties

For opportunity density $g(h,w)$, assume first that it is multiplicatively separable in h and w; that is, $g(h,w) = g_1(h)g_2(w)$, for $h > 0$, where $g_1(h)$ and $g_2(w)$ are normalized such that they become probability density functions. The function $g_1(h)$ is equivalent to the mean fraction of jobs with offered hours h that are feasible to the agent, and $g_2(w)$ is equivalent to the mean fraction of jobs with wage rates w that are feasible to the agent. See Dagsvik and Strøm (2004) for a justification of the separability assumption. Although offered wages and hours vary across jobs, this assumption implies that hours are set independent of wages. We shall discuss briefly the nonseparable case later.

Empirical evidence suggests that there seems to be a rather wide variation in the wage opportunities across agents, and it seems hardly possible to account for this distribution by observed individual characteristics such as observed length of schooling and potential experience. According to Mortensen (2003), observable worker characteristics that account for productivity differences typically explain no more than 30% of the variation in compensation across workers. The unexplained differences can be due to both unobserved individual abilities and unobserved job-specific effects. To account for unobserved individual-specific abilities in the wage offer distribution we introduce a random effect in this distribution. To this end, let $g_2(w \mid \eta)$ denote the conditional density of offered wage rates given η, where η is the individual specific random effect that is supposed to account for unobserved heterogeneity in wage rate opportunities (wage rate level) across the population. Recall that in the random choice set specification in Section 3.1 we introduced alternative specific random effects that were independent across alternatives and individuals. What we introduce here is an additional random effect η that is constant across different alternatives and solely individual specific.

Under the assumptions above it follows from (2.4) and (2.5) that the conditional choice probabilities given the random effect are equal to

(4.6) $$\varphi(h,w \mid I,\eta) = \frac{\theta\psi(h,w,I)g_1(h)g_2(w \mid \eta)}{\psi(0,0,I) + \theta \sum_{x\in D, x>0,\ y\in G} \sum \psi(x,y,I)g_1(x)g_2(y \mid \eta)},$$

for $h > 0$, and

(4.7) $$\varphi(0,0 \mid I,\eta) = \frac{\psi(0,0,I)}{\psi(0,0,I) + \theta \sum_{x\in D, x>0,\ y\in G} \sum \psi(x,y,I)g_1(x)g_2(y \mid \eta)},$$

for $h = 0$. The corresponding unconditional probabilities follow directly from (4.6) and (4.7). Unfortunately, the extension of the model leading to (4.6) and (4.7) raises formidable identification problems that cannot be easily solved. Essentially, in this extension the unexplained component in the wage rate is decomposed into a worker effect and a job-specific effect (including possible firm/industry effect). With crosssection data on labor supply typically available, this task is not possible. Several simplifications are used in the literature to cope with this problem.

A simple case is obtained when the wage rate does not vary across jobs for a given agent and the wage rate is observed (or depends only on observed individual characteristics). In this case, similar to the traditional neoclassical model of labor supply, the agent faces an individual-specific wage rate. It follows immediately from (4.6) and (4.7) that the conditional probability of choosing *h* hours of work given wage rate *w* in this case is given by

(4.8)
$$\varphi(h \mid w, I) = \frac{\theta \psi(h, w, I) g_1(h)}{\psi(0,0,I) + \theta \sum_{x \in D, x>0,} \psi(x, w, I) g_1(x)},$$

for $h > 0$, and similarly for $h = 0$. This simplification is empirically more tractable compared with (4.6) and (4.7). However, the assumption of constant wage rates across jobs is quite restrictive and often criticized.

Returning to the general case above, we shall now make an additional assumption, namely that the random effect η enters the offerd wage rate density solely through the mean, and as a positive multiplicative term. Let $\bar{w}\eta$ denote the mean of the conditional density $g_2(w \mid \eta)$. By Taylor expansion

(4.9)
$$\begin{aligned}\psi\left(h, W(z), I\right) &= \psi\left(h, \bar{w}\eta, I\right) + \left(W(z) - \bar{w}\eta\right)\psi_2'\left(h, \bar{w}\eta, I\right) \\ &\quad + 0.5\left(W(z) - \bar{w}\eta\right)^2 \psi_2'\left(h, \zeta\left(\bar{w}\eta, W(z)\right), I\right),\end{aligned}$$

where ζ lies between $\bar{w}\eta$ and $W(z)$ and ψ_2' denotes the partial derivative with respect to the second argument. It follows from (4.8) that

(4.10)
$$\sum_{y>0} \psi(h, y, I) g_2(y \mid \eta) = E_W\left(\psi(h, W(z), I) \mid \eta\right) \cong \psi(h, \bar{w}\eta, I),$$

where E_W denotes the expectation operator of the wage opportunity density. From (4.9) we realize that the approximation in (4.10) is close provided the variance in the wage opportunity distribution (conditional on the random effect) is small. By using (4.10), we can express the marginal conditional choice probability for hours of work given the random effect as

(4.11)
$$\varphi(h \mid I, \eta) = \sum_{w>0} \varphi(h, w \mid I, \eta) \cong \frac{\theta \psi(h, \bar{w}\eta, I) g_1(h)}{\psi(0,0,I) + \theta \sum_{x \in D, x>0,} \psi(x, \bar{w}\eta, I) g_1(x)},$$

for $h > 0$, and for $h = 0$ we obtain similarly that

(4.12) $$\varphi(0 \mid I,\eta) \cong \frac{\psi(0,0,I)}{\psi(0,0,I)+\theta \sum_{x\in D, x>0} \psi(x,\overline{w}\eta,I) g_1(x)}.$$

The unconditional marginal choice probability of hours of work is of course equal to

(4.13) $$\varphi(h \mid I) = E_\eta \varphi(h \mid I,\eta),$$

where E_η denotes the expectation operator with respect to η. One important implication of the random effect specification above is that the resulting choice probabilities will no longer satisfy the IIA assumption. This is the approach which is used in Dagsvik and Strøm (2006), Dagsvik and Jia (2006) and Kornstad and Thoresen (2007). Another variety is to ignore the unobserved individual effect and attribute the unobserved wage differential solely to job-specific variations. Aaberge, Dagsvik and Strøm(1995), Aaberge, Colombino and Strøm (1999) applied such a strategy. However, it remains an open question wich strategy is better in practice. A thorough empirical evaluation of these two approaches is needed to answer this question, since both are only approximations of the 'true' model (4.6)-(4.7).

One possibility which we intend to pursue in the future is to apply panel data to identify the wage offer distribution.[7] More precisely, suppose one is willing to assume the following representation of $g_2(w \mid \eta)$, expressed as

(4.14) $$\log W_t(z) = X_t\beta + \eta + \xi_t(z),$$

where t index time/age, X_t is a vector of suitable covariates and $\xi_t(z)$ is a job specific random term that is serially uncorrelated (or more generally following a suitable ARMA structure) and uncorrelated with the random effect η. Together with suitable distributional assumptions of η and $\xi_t(z)$, (4.14) implies is a particular structure of $g_2(w \mid \eta)$, which subsequently can be inserted into (4.6) and (4.7) with the purpose of maximum likelihood estimation based on panel data on chosen hours of work and wage rate combinations. We plan to address this issue in future research.

In the empirical version the choice probabilities in (4.11) and (4.12) also depend on socio-demographic variables that affect the systematic term of the utility function, $v(C,h)$. It is evident that without further assumptions on $v(C,h)$ and $g_1(h)$, one cannot separate $v(C,h)$ from $g_1(h)$. However, if

---

[7] This is closely related to the study of wage dispersion. See for example Mortenson (2003) for detailed discussion. In particular , we refer to Abowd, Kramarz and Margolis (1999) and Abowd and Kramarz (2001) for discussion on the decomposition of the wage dispersion into worker and employer components.

[7] Dagsvik and Strøm (2006) explicitly model the joint choice over sector as well as hours of work, in contrast to this paper in which we are only concerned with total labor supply, irrespective of hours of work.

the purpose of the model is to provide policy simulation results from changes in wage rates and the tax system (conditional on the choice sets) then there is no need to separate $g_1(h)$ from $v(C,h)$ because the effects from taxes and wage rates operate solely through the term $v\big(f(hw,I),h\big)$. Thus, as long as the product $v(C,h)g_1(h)$ is given an empirical representation in the model and θ is kept constant (since it is part of the choice set representation) there is no problem. In contrast, if one wishes to assess the effect of a reform that consists in changing the full time hours of work regulation then it is necessary to separate $v(C,h)$ from $g_1(h)$. In Aaberge, Dagsvik and Strøm (1995), Aaberge, Colombino and Strøm (1999), Dagsvik and Strøm (2006), Dagsvik and Jia (2006) it is assumed that $g_1(h)$ is uniform apart from peaks at full time- and part time hours. Analogous assumptions are made in the related labor supply literature that deals with rationing of hours, see for example Dickens and Lundberg (1993), Tummers and Woittiez (1991) and Bloemen(2000).

Consider next the more general case where the hours of work and wage rate offers are correlated, namely the opportunity measure is not separable. Barzel (1973) argued that wage offers may systematically vary hours of work, typically an inverted U-shaped relationship. (See also Wolf, 2002 for a collection of arguments.) This feature can be easily incorporated into our framework by specifying the density of offered wage rates depends on offered hours, that is $g_2(w\,|\,\eta)=g_2(w\,|\,h,\eta)$. This specification corresponds to analogous assumptions found in the literature, see for example Moffitt (1984), Tummers and Woittiez (1991) and Bloemen (2000). In the absence of random effects Dagsvik and Strom (1997) demonstrate that one can still identify nonparametrically the conditional offered distribution of wage rates, given offered hours of work, provided the deterministic part of the utility function is separable in disposable income and hours of work (leisure). The intuition is that whereas the utility of consumption component depends on disposable income through hours of work, wage rate and nonlabor income, the conditional density of offered wage rates depends only on the wage rate and hours of work. In the presence of random effect identification becomes much more problematic.

4.4. Aggregation of sectors

In this section we shall demonstrate that the modeling framework discussed above has some convenient aggregation properties. We have already showed (Section 3.1) that the model structure is invariant under aggregation of latent and heterogeneous choice sets. We shall now consider the case in which there are several latent sectors over which the agents have preferences[8]. Let $U(C,h,s,z)$ be the

[8] Dagsvik and Strøm (2006) explicitly model the joint choice over sector as well as hours of work, in contrast to this paper in which we are only concerned with total labor supply, irrespective of hours of work.

agent's utility of consumption, hours of work and job z in sector s. For simplicity, we assume that there are only two sectors, s = 1, 2, because the development in the general is completely analogous. We now extend (2.1) to

$$U(C,h,s,z) = v(C,h)b_s\varepsilon_s(z), \tag{4.15}$$

where b_s is a positive deterministic term that represents the mean preferences for working in sector s, whereas $\varepsilon_s(z)$ is, as above, a positive random term, possibly sector specific, with the same distributional assumptions as before. What essential makes the current setting different from the above one is that preferences over jobs in different sectors are, on average, different, and this is captured by the terms $\{b_s\}$. In addition, we allow the opportunity sets of jobs to differ across jobs. Consistent with the notation above, let $g_{1s}(h)$ denote the opportunity density of hours, given sector s. That is, $g_{1s}(h)$ represents the distribution of offered hours in sector s. Similarly to θ defined above, let θ_s represent the "size" of the job opportunity set within sector s. Let $\varphi(h,s\,|\,I)$ denote the joint probability of working in sector s with h hours of work. Thus the choice probability above has the interpretation as

$$\varphi(h\,|\,I) = \varphi(h,1\,|\,I) + \varphi(h,2\,|\,I)\,. \tag{4.16}$$

In Appendix D, we demonstrate that the assumptions above imply that the structure of $\varphi(h\,|\,I)$ still remains exactly as given above, but where now the opportunity density $g_1(h)$ and θ are to be interpreted as

$$g_1(h) = \frac{\theta_1 b_1 g_{11}(h) + \theta_2 b_2 g_{12}(h)}{\theta_1 b_1 + \theta_2 b_2} \text{ and } \theta = \theta_1 b_1 + \theta_2 b_2\,. \tag{4.17}$$

From (4.17) we see that if $b_1 \neq b_2$, and $g_{11}(h) \neq g_{12}(h)$, then the opportunity density of hours will depend on preferences through $\{b_s\}$, in contrast to the case with only one sector. Suppose for example that sector one contains more full time jobs than jobs with other hours whereas sector two contains more part time jobs than jobs with other hours. Then $g_{11}(h)$ will be a uniform density apart from a peak at full time hours whereas $g_{12}(h)$ will be uniform apart from a peak at part time hours. Suppose furthermore that $b_s, s = 1,2$, differ by gender. Then it follows from (4.17) that $g_1(h)$ differs by gender and is uniform apart from peaks at part time- and full time hours of work. In particular, the sizes of the peaks will differ by gender.

5. Relationship to other approaches

5.1. Comparison with the standard approach

In the standard approach to labor supply modeling, prior to the Hausman methodology, the typical approach was to choose a specification of an individual labor supply function (hours of work function) consistent with the maximization of a quasi-concave utility function in disposable income and leisure, subject to the economic budget constraint. For simplicity, consider the case with convex budget sets with constraint approximated by a suitable smooth representation. Suppose for example that the chosen labor supply function has the structure

(5.1) $$h = \alpha + \beta \tilde{w}(h) + X\gamma + \delta \tilde{I}(h) + \varepsilon,$$

when

(5.2) $$\alpha + \beta \tilde{w}(0) + X\gamma + \delta \tilde{I}(0) + \varepsilon > 0,$$

and $h = 0$ otherwise, where $\tilde{w}(h)$ is the marginal wage rate, $\tilde{I}(h)$ is so-called virtual nonlabor income, X is a vector of individual characteristics that affect preferences, ε is a random error term that is generated by (for example) a normal distribution $N(0,\sigma)$ and α, β, γ, σ and δ are unknown parameters. The inequality in (5.2) represents the condition for working. In general, when the tax system is nonlinear, the marginal wage rate and virtual income depend on hours of work and hence, they are endogenous. As a result, one cannot estimate the model by using OLS. Additional complications follow from the condition given by (5.2) and from the fact that the wage rate is not observed for those who do not work. Now, suppose that the parameters of this labor supply function have been estimated. Then, to derive the hours of work relation when (5.2) holds, given the wage rate and nonlabor income, one needs to solve for h in the nonlinear equations given by (5.1) and (5.2). Let us denote by $h_i = F\left(w_i, I_i, X_i, \varepsilon_i\right)$ the resulting labor supply function of worker i, obtained by solving for hours of work, where w_i is the wage rate and I_i is nonlabor income for worker i. Then, one can simulate the conditional distribution of labor supply (hours of work), given (w_i, I_i, X_i), by drawing T i.i.d. error terms $\{\varepsilon_i\}$ from the normal distribution $N(0,\sigma)$, and compute the this conditional distribution as

(5.3) $$P\left(h_i \le y \mid w_i, I_i, X_i\right) = \sum_{\{k; F(w_i, I_i, X_i, \varepsilon_k) \le y\}} F\left(w_i, I_i, X_i, \varepsilon_k\right) \frac{1}{T}.$$

The summation on the right-hand side of (5.3) is taken over all k such that supplied hours are less than or equal to y. The empirical counterpart of (5.3) is the fraction of agents with characteristics (w_i, I_i, X_i) that supply hours of work of less than or equal to y. The corresponding unconditional labor supply distribution can be obtained by computing

(5.4) $$P(h_i \le y) = \frac{1}{N} \sum_{i \in \Omega} P\left(\tilde{h}_i \le y \mid w_i, I_i, X_i\right),$$

where Ω denotes a representative micro-population of size N. In principle, this can be done with general utility specifications and the corresponding labor supply functions, but, as mentioned above, this will in most cases be rather cumbersome in practice. The reason for this is that the class of utility functions that imply explicit labor supply functions, such as the one in (5.1), is rather limited and, hence, in more general cases, one is forced to work with nonlinear specifications. Thus, even when the budget constraint is simplified to ensure a convex budget set, the estimation and simulation of labor supply responses is not straightforward.

In contrast, the alternative modeling framework proposed in this paper offers several advantages compared to the standard approach. First, it allows for preferences to depend on nonpecuniary job attributes, and second, it accounts for possible constraints on the set of feasible job offers. In addition to these theoretical aspects, this framework does not require an explicit derivation of an individual labor supply function. Instead, the distribution (probability density) of labor supply is modeled directly and expressed explicitly as a function of the systematic term of the utility function. As a result, one needs not simplify the budget constraint in (2.2). In addition, the systematic term of the utility function, $v(C, h)$, can be quite general because this approach does not depend on the derivations of, and solutions to, first-order conditions. Furthermore, the simulation of distributional effects is straightforward because, as mentioned above, the model is represented explicitly in terms of a probability density, cf. (4.11) and (4.12).

5.2. Comparison with the conventional discrete choice approach

The conventional discrete choice approach also shares many of the practical features with our approach in that no marginal calculation is needed, cf. van Soest (1995). See also the review by Creedy and Kalb (2005). Specifically, it enables the researcher to straightforwardly apply quite general specifications of the utility function and the budget constraint. However, as mentioned in the introduction, it does not accommodate the feature that preferences typically depend on nonpecuniary job attributes, and because it is basically a version of the standard approach, it cannot accommodate peaks on full-time and part-time hours nor restrictions on the set of feasible jobs. Thus, under the analogous assumption to (2.1), and with extreme value distributed error terms the choice probability $\tilde{\varphi}(h \mid w, I)$ (say) of the standard discrete choice model that corresponds to (2.4) takes the form

$$\tilde{\varphi}(h \mid w, I) = \frac{\psi(h, w, I)}{\psi(0,0,I) + \sum_{x \in D, x > 0} \psi(x, w, I)}, \tag{5.9}$$

which we realize can be viewed as a special case that follows from (2.4) when the opportunity measure $\theta g(h,w)$ is independent of h and equal to zero for w different from a given individual-specific wage rate and equal to one otherwise. In other words, we note that the standard discrete choice approach is a special case of ours. Evidently, with conventional assumptions about preferences in the model in (5.9) one cannot without further assumptions explain the high concentrations on hours of work at full time hours, and also possibly at part time hours, a feature which is typical in many countries. This is a consequence of the fact that the conventional discrete choice labor supply model is based on the standard theoretical framework in which the agent is assumed to have preferences over leisure and disposable income combinations only, and is able to freely choose leisure and disposable income combinations subject to the budget constraint and discrete hours.

Van Soest (1995) and other reseachers estimate a model of type (5.9) and note that his model is unable to fit his data due to a peak at full time hours. Typically, they subsequently extend the model by including additional alternative specific constants for alternatives that are not full-time alternatives. The argument is that these constants reflect choice restrictions that operate through monetary or non-monetary drawbacks of not working full time. This procedure is however, purely ad hoc and unsatisfactory because the modeling framework represented by (5.9) does not allow for this type of choice restrictions due to the fact that no explicit rationing device or quantity restrictions follow from the theory, apart from the assumption that the choice set of hours is discrete, and possibly a parameter that represents fixed cost of work. In contrast, our alternative approach advocated in this paper provides a theoretical justification for accommodating restrictions on hours of work by allowing the opportunity density of hours of work to have peaks at full-time and possibly part time hours of work, and the parameter θ to account for the degree of job availability. In fact, the parameter θ can, alternatively, also be interpreted as due to fixed cost of working (latent), or rather due to a combination of fixed cost of working and degree of job availability. As will be discussed further in Section 5.3, this type of specification is mathematically equivalent to introducing suitable dummy variables in the utility specification of the model in (5.9).

5.3. Extended versions of the Hausman approach with continuous choice

First interpretation:

Evidently, the discrete choice approach of van Soest (1995) with suitable random effects can be viewed as a version of the Hausman approach extended to allow for a separable extreme value distributed error term that vary across hours of work and restricted to a finite set of feasible hours. We shall now demonstrate that a similar continuous choice model also exists. To this end, assume for expository simplicity and in complete accordance with the assumptions made in the Hausman type of models, that the wage rate is fixed for a given individual and that the opportunity distribution of hours is uniform. Suppose furthermore that the coefficients of $\psi(h,w,I)$ are random. Assume now that D is a continuous set and let B be the countable subset of D consisting of an infinite number of independent and uniformly distributed hours $H(z), z=1,2,...,$ on D. Since the number of random draws in B are infinite it follows that B is a countable dense subset of D. Thus, there is no essential difference between B and D as representations of the set of feasible hours. Furthermore, let $\{\tilde{\varepsilon}(H(z)), z=1,2,...,\}$ be an enumeration of points of a Poisson process on $(0,\infty)$ with intensity measure $\varepsilon^{-2}d\varepsilon,$ which are independent of B. Then by Resnick (1987, p.135) it follows that the points $\{(H(z),\tilde{\varepsilon}(H(z)), z=1,2,...\},$ are the points of a Poisson process on $D\times(0,\infty),$ with intensity measure $\varepsilon^{-2}d\varepsilon dh.$ Note that the formalism here is slightly different from the one above in that here it is convenient to let $H(z)$ represent the indexation. Apart from interpretation, this represents no loss of generality. Let $\tilde{U}(h)$ be the conditional indirect utility, given hours of work, which we assume has the structure

$$\text{(5.10)} \qquad \tilde{U}(h)=\psi(h,w,I)\tilde{\varepsilon}(h)^{\sigma},$$

where $\sigma>0$ is a scalar. Let $\varphi^{*}(h\mid w,I)$ be the probability that the agent chooses hours of work h, given that the choice is determined by maximizing the utility function in (5.10) subject to hours of work belonging to the set B. Similarly to (3.10) it follows that

$$\text{(5.11)} \qquad \varphi^{*}(h\mid w,I)=E\left\{\frac{\psi(h,w,I)^{1/\sigma}}{\psi(0,0,I)^{1/\sigma}+\int\limits_{x\in D,x>0}\psi(x,w,I)^{1/\sigma}dx}\right\},$$

where the expectation operator now is taken with respect to the random coefficients of the utility function. The proof of (5.11) is completely similar to the proof of (3.10). Moreover, if $\sigma\to 0,$ then by (5.10) $\tilde{U}(h)\to\psi(h,w,I)$, in which case the utility maximization problem reduces to the maximization of $v(C,h)$ subject to the budget constraint. Specifically, in the typical version of the Hausman approach the utility function $v(C,h)$ is specified as

(5.12) $$v(C,h) = \left(\frac{h-\alpha_1}{\alpha_2}\right)\exp\left(\frac{\alpha_2\left(C+\alpha_3\right)}{h-\alpha_1}\right),$$

where α_1 and α_3 are unknown parameters and $\alpha_2 \leq 0,$ is a random coefficient. Thus, the only difference between the specification given in (5.10) and (5.11), and the standard Hausman model is the assumptions about the additional separable stochastic term $\tilde{\varepsilon}\left(h\right)^{\sigma}$ in the utility function. Whereas the standard econometric technique for estimating the Hausman model (with piecewise non-convex budget sets) is very complicated the extended version given in (5.10) and (5.11) appears to be much simpler.

Second interpretation:

As discussed in the previous section the discrete approach of van Soest can also be viewed as a special case of our framework (and accordingly the Hausman approach) with choice or job as the essential choice variable. This follows when the opportunity measure representing restrictions on hours of work is assumed to be uniform. The corresponding utility of hours *h*, $U^*(h)$, has in this case the interpretation as an indirect utility function given hours of work, and is expressed as

$$U^*(h) = \psi(h,w,I)\max_{z\in B(h)}\varepsilon(z),$$

where $B(h)$ is the set of feasible jobs with offered hours of work equal to *h*. As discussed in Section 3.1 there is also a continuous choice version of the job choice model with corresponding choice probability given in (3.10). From (3.10) it follows that when the wage rate is fixed and the opportunity measure is uniform the expression in (3.10) reduces to the one in (5.11).

6. Conclusion

In this paper, we have focused on what we believe are some key issues in empirical modeling of structural labor supply relations. In particular, we have argued for the application of a particular alternative modeling approach for empirical analysis of labor supply behavior that differs from conventional ones. We have contrasted our framework with standard approaches found in the literature, such as the Hausman approach and the conventional discrete choice approach, and various approaches aiming at taking choice restrictions into account. An essential feature of our framework is that it is consistent with the notion of latent job opportunities, which we have shown imply that one can easily accommodate restrictions on the set of feasible jobs, and peaks in the hours of work distribution (interpreted as due to restrictions on hours), typically observed in many data sets. A substantial part of the paper is addressed to extensions and further interpretation of the framework, such as unobserved heterogeneity in latent choice sets, identification and functional form. In particular, we demonstrate that several approaches in the literature can be viewed as embedded in our approach.

An important point is that the alternative approach we propose is *practical* for empirical analysis and policy simulation experiments. This has been demonstrated in empirical applications reported in Dagsvik and Jia (2006), Dagsvik and Strøm (2006).

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.

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.

Abowd, J. and F. Kramarz (2001): Inter-industry and Firm-size Wage Differentials: New Evidence from Linked Employer-Employee Data. Paper presented at the Econometric Society World Conference, 2000.

Abowd, J., F. Kramarz and D. Margolis (1999): High Wage Workers and High Wage Firms. *Econometrica* **67**, 251–334.

Barzel, Y. (1973): The Determination of Daily Hours and Wages. *Quarterly Journal of Economics,* **87**, 220–238.

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

Ben-Akiva, M., N. Litinas and K. Tsunokawa (1985): Spatial Choice: The Continuous Logit-Model and Distribution of Trips and Urban Densities. *Transportation Research Part* A, *19A*, 119-154.

Bingley, P. and I. Walker (1997): The Labour Supply, Unemployment and Participation of Lone Mothers in In-work Transfers Programmes. *Economic Journal*, **107**, 1375-1390.

Bloemen, H. (2000): A model of labor supply with job offer restrictions. *Labour Economics*, **7**, 297-312.

Bloemen, H. (2008): Job Search, Hours Restrictions, and Desired Hours of Work. *Journal of Labor Economics*, **26**, 137-179.

Bloemen, H. and A. Kapteyn (2008): The estimation of utility consistent labor supply models by means of simulated scores. *Journal of Applied Econometrics*, **23**, 395-422.

Blomquist, S. (1983): The Effect of Income Taxation on the Labor supply of Married Men in Sweden. *Journal of Public Economics*, **17**, 169-197.

Blomquist, S. (1992): Estimation Methods for Male Labor Supply Functions. How to Account for Nonlinear Taxes. *Journal of Econometrics,* **70**, 383-405.

Blundell, R. W., J. Ham and C. Meghir (1987): Unemployment and Female Labour Supply. *Economic Journal,* **97**, 44-64.

Blundell, R. W. and T. MaCurdy (1999): Labor Supply: A Review of Alternative Approaches. In O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Vol. 3A. North-Holland, Amsterdam.

Blundell, R., A. Duncan, J. McCrae and C. Meghir (2000): The Labour Market Impact of the Working Families' Tax Credit. *Fiscal Studies,* **21**, 75-104.

Burtless, G. and J. A. Hausman (1978): The Effects of Taxation on Labor Supply. *Journal of Political Economy*, **86**, 1103-1130.

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

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

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

Dagsvik, J. K. (2008): Discrete Choice and Aggregation of Latent Choice Sets. Manuscript, Statistics Norway.

Dagsvik, J. K. and Z. Jia (2006): Labor Supply as a Choice among Latent Choice Opportunites. A Practical Empirical Approach. Discussion Paper no. 481, Statistics Norway.

Dagsvik, J. K. and S. Røine (2008): Testing Qualitative Functional Form Assumptions in Nonlinear Models. Manuscript, Statistics Norway.

Dagsvik, J. K. and S. Strøm (1988): A Labor Supply Model for Married Couples with Non-convex Budget Sets and Latent Rationing. Discussion Paper no. 36, Statistics Norway.

Dagsvik, J. K. and S. Strøm (1997): A Framework for Labor Supply Analysis in the Presence of Complicated Budget Restrictions and Qualitative Opportunity Aspects. Memorandum no. 22, Department of Economics, University of Oslo.

Dagsvik, J. K. and S. Strøm (2004): Sectoral Labor Supply, Choice Restrictions and Functional Form. Discussion Paper no. 388, Statistics Norway, Oslo.

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

Dagsvik, J. K., S. Strøm and Z. Jia (2006): Utility of Income as a Random Function: Behavioral Characterization and Empirical Evidence. *Mathematical Social Sciences*, **51**, 23–57.

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

Duncan, A. and C. Giles (1998): The Labour Market Impact of the Working Families' Tax Credit. Paper presented at the International Institute for Public Finance Conference, Cordoba, Spain.

Duncan, A. and M. Weeks (1998): Simulating Transitions using Discrete Choice Models. *Papers and Proceedings of the American Statistical Association*, **106**, 151-156.

Embrechts, P., C. Klüppelberg and T. Mikosch (1997): *Modelling Extremal Events*. Springer-Verlag, Berlin.

Falmagne, J. C. (1985): *Elements of Psychophysical Theory*. Oxford University Press, London.

Falmagne, J. C. and G. Iverson (1985): Statistical Issues in Measurement. *Mathematical Social Sciences*, **10**, 131-153.

French, E. (2005): The Effects of Health, Wealth, and Wages on Labor Supply and Retirement Behaviour. *Review of Economic Studies*, **72**, 395-427

Hausman, J. A. (1985): Taxes and Labor Suppy. In A. Auerback and M. Feldstein (eds.), *Handbook of Public Economics*, Vol I. North-Holland, Amsterdam.

Hausman, J. A. and P. Ruud (1984): Family Labor Supply with Taxes. *American Economic Review*, **74**, 242–253.

Heim, B. T. and B. D. Meyer (2003): Structural Labor Supply when Budget Constraints are Non-linear. Working Paper, The Harris School of Public Policy, University of Chicago.

Hoynes, H. W. (1996): Welfare Transfers in Two-Parent Families: Labor Supply and Welfare Participation under the AFDC-UP Program. *Econometrica*, **64**, 295-332.

Hyslop, D. (1999): State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labor Force Participation of Married Women, *Econometrica*, **67**, 1255–1294.

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

Kapteyn, A., P. Kooreman and A. van Soest (1990): Quantity Rationing and Concavity in a Flexible Household Labor Supply Model. *Review of Economics and Statistics*, **62**, 55–62.

Keane, M. and R. Moffitt (1998): A Structural Model of Multiple Welfare Program Participation and Labor Supply. *International Economic Review*, **9,** 553-589.

Keane, M. and K. Wolpin (2001): Estimating Welfare Effects Consistent with Forward-Looking Behavior: Part II: Empirical Results. *Journal of Human Resources*, **37**, 600–622.

Kornstad, T. and T. O. Thoresen (2007): A Discrete Choice Model for Labor Supply and Child Care. *Journal of Population Economics*, **20**, 781-803.

Luce, R. D. (1959): *Individual Choice Behavior.* Wiley, New York.

Luce, R. D. (1977): The Choice Axiom after Twenty Years. *Journal of Mathematical Psychology*, **15**, 215-233.

MaCurdy, T. E. (1992): Work Disincentive Effects of Taxes: A Re-examination of some Evidence. *American Economic Review*, **82**, 243-249.

MaCurdy, T. E., D. Green and H. Paarsch (1990): Assessing Empirical Approaches for Analyzing Taxes and Labor Supply. *Journal of Human Resources*, **25**, 415-490.

McFadden, D. (1984): Econometric Analysis of Qualitative Response Models. In Z. Griliches and M.D. Intriligator (eds.); *Handbook of Econometrics,* Vol. II, Elsevier Science Publishers BV, New York.

McFadden, D. and K. Train (2000): Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics*, **15**, 447-470.

Moffitt, R. (1984): The estimation of a joint wage-hours labor supply model. *Journal of Labor Economics,* **4**, 550–566.

Morrison, H. W. (1962): *Intransitivity of Paired Comparison Choice.* Unpublished doctoral dissertation, University of Michigan, Ann Arbor.

Mortensen, D. T. (2003): *Wage Dispersion: Why Are Similar Workers Paid Differently*? MIT Press, MA, USA.

Narens, L. (2002): *Theories of Meaningfulness.* Lawrence Erlbaum Associates, Publishers, Mahwah, New Jersey, USA.

Quandt, R. (1956): Probabilistic Theory of Consumer Behavior. *Quarterly Journal of Economics*, **70**, 507-536.

Resnick, S. I. (1987): Extreme Values, Regular Variation, and Point Processes. Springer-Verlag, Berlin.

Samorodnitsky, G. and M.S. Taqqu (1994): *Stable Non-Gaussian Random Processes.* Chapman & Hall, London.

Sattath, S. and A. Tversky (1976): Unite and Conquer: A Multiplicative Inequality for Choice Probabilities. *Econometrica*, **44**, 79-89.

Simon, H. (1986): Rationality in Psychology and Economics. *Journal of Business,* **59**, S209-224.

Stevens, S. S. (1975): *Psychophysics: Introduction to its Perceptual Neural, and Soical Prospects*. Wiley, New York.

Thurstone, L. L. (1927): A Law of Comparative Judgment. *Psychological Review*, **34**, 153-181.

Tummers, M. P. and I. Woittiez (1990): A simultaneous Wage and Labor Supply Models with Hours restrictions. *Journal of Human Resources*, **26**, 393-423.

Tversky, A. (1969): Intransitivity of Preferences. *Psychological Review*, **84**, 327-352.

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

van Soest, A., I. Woittiez and A. Kapteyn (1990): Labor Supply, Income Taxes, and Hours Restrictions in the Netherlands. *Journal of Human Resources* **25**, 517-558.

van Soest, A., M. Das and X. Gong (2002): A Structural Labour Supply Model with Flexible Preferences. *Journal of Econometrics*, **107**, 345–374.

Wolf, E. (2002): Lower Wage Rates for Fewer Hours? A Simultaneous Wage-hours Model for Germany. *Labour Economics*, **9**, 643-663.

Unobserved heterogeneity in choice sets as random effects

For expositional simplicity, we shall in the following simplify notation in proving (3.4), by dropping the wage rate *w* and nonlabor income *I* in the notation. This represents no essential loss of generality because the argument remains the same. Define the conditional choice probability given the choice sets (analogous to (3.2)) by

(A.1) $$\varphi\left(h \mid \{\mu(h'), h' \in D\}\right) = \frac{\psi(h)\mu(h)}{\sum_x \psi(x)\mu(x)}.$$

It follows from (A.1) that the set $\{\mu(h), h \in D\}$ represents a sufficient set of random variables for the latent choice sets $\{B(h), h \in D\}$. Write $\mu(h) = m(h)\omega(h)$. Recall the following assumptions: (i) $\omega(h) > 0$; (ii) for any hours of work, h_1 and h_2, and nonnegative constants, b_1 and b_2, then $b_1\omega(h_1) + b_2\omega(h_2)$ have the same distribution as $c\omega(h_1)$, where *c* is a positive constant that may depend on h_1, h_2, b_1 and b_2; (iii) the random variables $\omega(h_1)$ and $\omega(h_2)$ are i.i.d. These assumptions imply that, for any *h*, $\omega(h)$ is strictly stable with scale parameter that is independent of *h* and can therefore be normalized to 1. Moreover, it is totally skew to the right with index $\alpha < 1$. Let *A* be a subset of *D* and define

(A.2) $$\varphi\left(A \mid \{\omega(h'), h' \in D\}\right) \equiv \sum_{x \in A} \varphi\left(x \mid \{\omega(h'), h' \in D\}\right).$$

It follows from the properties of stable random variables that

(A.3) $$\sum_{x \in A} \psi(x) m(x) \omega(x) \overset{d}{=} \left(\sum_{x \in A} (\psi(x) m(x))^{\alpha} \right)^{1/\alpha} \xi(A),$$

where $\overset{d}{=}$ means equality in distribution and $\xi(A)$ is distributed according to $S_\alpha(1,1,0)$. This means that we can express the probability in (A.2) as

(A.4) $$\varphi\left(A \mid \{\omega(h'), h' \in D\}\right) \overset{d}{=} \frac{\left(\sum_{x \in A} (\psi(x) m(x))^{\alpha}\right)^{1/\alpha} \xi(A)}{\left(\sum_{x \in A} (\psi(x) m(x))^{\alpha}\right)^{1/\alpha} \xi(A) + \left(\sum_{x \in D \setminus A} (\psi(x) m(x))^{\alpha}\right)^{1/\alpha} \xi(D \setminus A)},$$

where $\xi(A)$ and $\xi(D\backslash A)$ are independent and distributed according to $S_\alpha(1,1,0)$. The structure of the formula in (A.4) implies that the distribution of $\varphi\left(A \mid \{\omega(h'), h' \in D\}\right)$ only depends on A through the deterministic parts

$$\left(\sum_{x\in A}\psi(x)^\alpha m(x)^\alpha\right)^{1/\alpha} \quad \text{and} \quad \left(\sum_{x\in D\backslash A}\psi(x)^\alpha m(x)^\alpha\right)^{1/\alpha},$$

since the distribution of ($\xi(A)$, $\xi(D\backslash A)$) does not depend on A nor on D. In other words, we have demonstrated the claim made in Section 3.1 that the structure of the distribution of the expression in (A.2) is invariant under aggregation of alternatives.

Next, define $U(h)=\psi(h)\mu(h)\varepsilon(h)=\psi(h)m(h)\omega(h)\varepsilon(h)$, for $h\in D$, where $\{\varepsilon(h)\}$ are i.i.d. positive random variables with c.d.f. $\exp(-1/x)$, for $x>0$. Conditional on $\{\mu(h)\}$ it follows that maximizing the random function $U(h)$ yields that the probability that the highest value occurs at h is equal to $\varphi\left(h \mid \{\omega(h'), h' \in D\}\right)$, given in (A.1). Suppose now that $\{\omega(h)\}$ are i.i.d. with distribution $S_\alpha(1,1,0)$. Consider the distribution of $\omega(h)\varepsilon(h)$. Since $\omega(h)$ and $\{\varepsilon(h)\}$ are independent, it follows from Proposition 1.2.12 in Samorodnitsky and Taqqu (1994, p. 15) that

$$\text{(A5)} \qquad P\left(\omega(h)\varepsilon(h)\le x\right)=EP\left(\varepsilon(h)\le\frac{x}{\omega(h)}\middle|\omega(h)\right)=E\exp\left(-\omega(h)x^{-1}\right)=\exp\left(-kx^{-\alpha}\right)$$

where $k=1/\cos(\alpha\pi/2)$. This implies that $\omega(h)\varepsilon(h)$ has the same distribution as $k^{1/\alpha}\varepsilon(h)^{1/\alpha}$. Consequently, this implies that the utility function $U(h)$ is equivalent to the utility function $\tilde{U}(h)=\psi(h)^\alpha m(h)^\alpha\varepsilon(h)$ because $k^{-1}U(h)^\alpha$ has the same distribution as $\tilde{U}(h)$. According to standard results in random utility theory, it follows that

$$\text{(A.6)} \qquad P\left(\tilde{U}(h)=\max_r \tilde{U}(r)\right)=\frac{\psi(h)^\alpha m(h)^\alpha}{\sum_{r\in D}\psi(r)^\alpha m(r)^\alpha},$$

which implies that

$$\text{(A.7)} \qquad \begin{aligned} E\left(\frac{\psi(h)m(h)\omega(h)}{\sum_{r\in D}\psi(r)m(r)\omega(h)}\right) &= EP\left(U(h)=\max_r U(r)\middle|\{\omega(r)\}\right) \\ &= P\left(\tilde{U}(h)=\max_r \tilde{U}(r)\right)=\frac{\psi(h)^\alpha m(h)^\alpha}{\sum_{r\in D}\psi(r)^\alpha m(r)^\alpha}. \end{aligned}$$

Q.E.D.

Random choice sets generated by a Poisson process

By (2.1) and (2.2),

(B.1) $$U(z)=\psi(H(z),W(z))\varepsilon(z),$$

where we for simplicity have suppressed non-labor income *I* in the notation.

Recall that $\{(H(z),W(z),\varepsilon(z)),z=1,2,...\}$ are realizations of a Poisson process on $D\times G\times R_{+}$ with intensity measure $d\zeta(h,w,\varepsilon)$ defined by

(B.2) $$d\zeta(h,w,\varepsilon)=m(h,w)\lambda(\varepsilon)dh\,dw\,d\varepsilon$$

for $h>0,\ w>0,\ \varepsilon>0$. Let *A* be a Borel set in $D\times G$, and define

(B.3) $$U_A=\max_z\{(\psi(H(z),W(z)))\text{ subject to }(H(z),W(z))\in A\}.$$

Thus U_A is the highest utility the agent can attain, subject to $(H(z),W(z))\in A$. Let

$$C_A(u)=\{(h,w,\varepsilon):\psi(h,w)\in A,\psi(h,w)\varepsilon>u\}$$

and let *N*(*C*) be the number of Poisson process points within *C*. By the Poisson law

(B.4) $$P(N(C_A(u))=n)=\frac{\Delta(C_A(u))^n}{n!}\exp(-\Delta(C_A(u))),$$

where $\Delta(C)=EN(C)$, and is given by

(B.5) $$\Delta(C_A(u))=\int_{C_A(u)}d\zeta(x,y,\varepsilon)=\int_{(x,y)\in A,}\int_{\psi(x,y)\varepsilon>u}\lambda(\varepsilon)m(x,y)d\varepsilon\,dx\,dy=\int_A\Lambda\left(\frac{\psi(x,y)}{u}\right)m(x,y)dx\,dy$$

where

$$\Lambda(x)=\int_{1/x}^{\infty}\lambda(\varepsilon)d\varepsilon.$$

This is so because $d\zeta(h,w,\varepsilon)$ is the probability that there is a Poisson point within $(h,h+dh)\times(w,w+dw)\times(\varepsilon,\varepsilon+d\varepsilon)$ and $\int_{C_A(u)} d\zeta(x,y,\varepsilon)$ is the "sum" of probabilities of all possible Poisson point locations within $C_A(u)$. It follows from (B.4) and (B.5) that

(B.6)
$$\begin{aligned} P(U_A \le u) &= P(\text{There are no points of the Poisson process in } C_A(u)) \\ &= P(N(C_A(u))=0) = \exp(-\Delta(C_A(u))) = \exp\left(-\int_A \Lambda\left(\frac{\psi(x,y)}{u}\right) m(x,y)\,dx\,dy\right) \end{aligned}$$

Since the points of the Poisson process are independent it follows that for two disjoint Borel sets *A* and *B* in $D\times G$, U_A and U_B are independent. Hence,

(B.7)
$$\begin{aligned} P(U_A \ge U_B) &= \int_0^\infty P(U_A \in du) P(u \ge U_B) \\ &= \exp\left(-\int_{A\cup B} \Lambda\left(\frac{\psi(x,y)}{u}\right) m(x,y)dxdy\right) \int_A \Lambda'\left(\frac{\psi(x,y)}{u}\right) \frac{m(x,y)dxdy}{u^2} du. \end{aligned}$$

The last expression can be rewritten as

(B.8)
$$\begin{aligned} &= \exp\left(-\int_{A\cup B} \Lambda\left(\frac{\psi(x,y)}{u}\right) m(x,y)dxdy\right) \int_{A\cup B} \Lambda'\left(\frac{\psi(x,y)}{u}\right) \frac{\psi(x,y)m(x,y)dxdy}{u^2} \\ &\times \frac{\int_A \Lambda'\left(\frac{\psi(x,y)}{u}\right)\psi(x,y)m(x,y)dxdy}{\int_{A\cup B} \Lambda'\left(\frac{\psi(x,y)}{u}\right)\psi(x,y)m(x,y)dxdy} du = E_U\left\{\frac{\int_A \Lambda'\left(\frac{\psi(x,y)}{u}\right)\psi(x,y)m(x,y)dxdy}{\int_{A\cup B} \Lambda'\left(\frac{\psi(x,y)}{u}\right)\psi(x,y)m(x,y)dxdy}\right\}. \end{aligned}$$

Let $A=(x,x+dx)\times(y,y+dy)$ and $B=D$. Then (3.7), (3.8) and (3.9) follow.

Q.E.D.

A Nested Multinomial Logit model version

Let the Utility function be defined as in (2.1) and let

(C.1)
$$\tilde{U}(h)=\psi(h,w,I)\tilde{\varepsilon}(h)=\psi(h,w,I)\max_{z\in B(h,w)}\varepsilon(z).$$

The interpretation of $\tilde{U}(h)$ is as the utility of the most preferred among the available jobs with hours of work h. Let B be the set of available jobs and $B(h)$ the set of available jobs with hours of work h. Assume that

(C.2)
$$P\left(\bigcap_{z\in B}(\varepsilon(z)\le u_z)\right)=\exp\left(-u_0^{-1}-\left(\sum_{z\in B,z>0}u_z^{-1/\rho}\right)^{\rho}\right),$$

which implies that

(C.3)
$$P\left(\bigcap_{h\in D}(\tilde{\varepsilon}(h)\le u_h)\right)=P\left(\bigcap_{h\in D}\left(\bigcap_{z\in B(h)}(\varepsilon(z)\le u_h)\right)\right)=\exp\left(-u_0^{-1}-\left(\sum_{h\in D,h>0}\theta g_1(h)u_h^{-1/\rho}\right)^{\rho}\right).$$

The interpretation of the parameter ρ is as

$$Corr\left(\varepsilon(z),\varepsilon(z')\right)=1-\rho^2,\text{ for }z\neq z',z>0,z'>0,\ Corr\left(\varepsilon(z),\varepsilon(0)\right)=0,\ z>0,$$

cf. McFadden (1984). We can now apply the results of McFadden (1984) that yields the choice probabilities for Generalized Extreme Value models, from which (4.13) follows.

Q.E.D.

Aggregation of sectors

By straightforward extension of the argument in Sections 2 and 4.3, it follows that similarly to (4.10), we obtain that

(D.1) $$\varphi(h,s\,|\,I) \cong E\left(\frac{\psi\left(h,\bar{w}\eta,I\right)\theta_s b_s g_{1s}(h)}{\psi(0,0,I)+\sum\limits_{x\in D,x>0,}\left[\psi\left(x,\bar{w}\eta,I\right)\sum\limits_r \theta_r b_r g_{1r}(x)\right]}\right),$$

where θ_s, b_s and $g_{1s}(h)$ are defined in Section 4.3. From (4.15) it follows immediately that

(D.2) $$\varphi(h\,|\,I) = \varphi(h,1\,|\,I) + \varphi(h,2\,|\,I) \cong E\left(\frac{\psi(h,\bar{w}\eta,I)\theta g_1(h)}{\psi(0,0,I)+\theta\sum\limits_{x\in D,x>0,}\psi(x,\bar{w}\eta,I)g_1(x)}\right),$$

where

(D.3) $$g_1(h) = \frac{\theta_1 b_1 g_{11}(h) + \theta_2 b_2 g_{12}(h)}{\theta_1 b_1 + \theta_2 b_2} \quad \text{and} \quad \theta = \theta_1 b_1 + \theta_2 b_2.$$

The relations in (D.3) are equal to (4.17) in sector 4.3.

Q.E.D.