

Statistics Norway  
Research Department

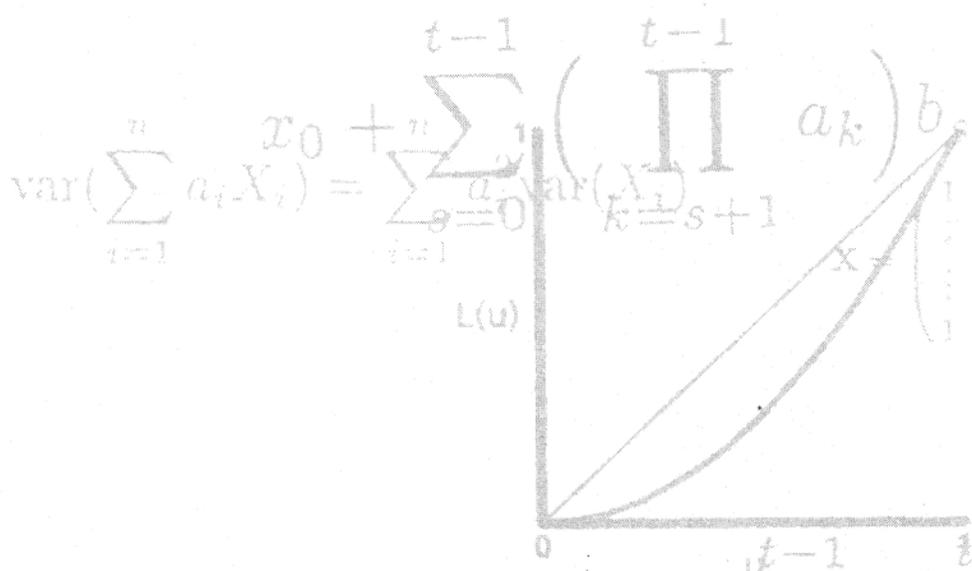
John K. Dagsvik

### Dynamic Choice, Multistate Duration Models and Stochastic Structure

# Discussion Papers

$$+ 2 \sum_{i>j} \sum_{j=1} \text{COV}_{at}(X_i, X_j)$$

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}$$



$$\text{var}(\sum_{i=1}^n a_i X_i) = \sum_{i=1}^n a_i^2 \text{var}(X_i) + \sum_{k=s+1}^{t-1} (\prod_{k=s+1}^{t-1} a_k) \sum_{i=1}^n (y_i - (\hat{a}x_i + \hat{b}))$$

*John K. Dagsvik*

## **Dynamic Choice, Multistate Duration Models and Stochastic Structure**

**Abstract:**

An important problem in the analysis of intertemporal choice processes is how to justify the choice of mathematical structure of the transition probabilities. A related and delicate identification problem is to separate the effect of unobserved variables from the influence on preferences from past choice behavior (state dependence). The present paper proposes a particular behavioral assumption to characterize the stochastic structure of intertemporal discrete choice models under the absence of state dependence. This assumption extends Luce axiom; "Independence from Irrelevant Alternatives", to the intertemporal context. Under specific regularity conditions the implication of these assumptions is that the individual choice process is a Markov chain with a particularly simple structure of the transition probabilities. By drawing on results obtained by Dagsvik (1983, 1988) it is demonstrated that this structure is consistent with an intertemporal and life cycle consistent random utility model where the utilities are independent extremal processes in time. Finally, the framework is extended to allow for state dependence and time varying choice sets.

**Keywords:** Life cycle consistent discrete choice, taste persistence, state dependence, Markovian choice processes, extremal processes, random utility processes, independence from irrelevant alternatives.

**JEL classification:** C25, C41, D91

**Acknowledgement:** Thanks to Rolf Aaberge, Tom Kornstad, Pedro de Lima, Chris Flinn and Steinar Strøm for criticism and comments and Anne Skoglund for word processing and correction of errors. I also thank participants in workshop at the University of California, Berkeley, University of Chicago, Université de Toulouse and University of Wisconsin.

**Address:** John K. Dagsvik, Statistics Norway, Research Department,  
P.O.Box 8131 Dep., N-0033 Oslo, Norway. E-mail: [jda@ssb.no](mailto:jda@ssb.no)

## 1. Introduction

In the literature on econometric analyses of duration data it is common practice to postulate the proportional hazard rate framework because it is convenient for empirical analyses. Needless to say, specifications based solely on mathematical convenience is ad hoc from a theoretical point of view and is therefore unsatisfactory. A theoretical justification supporting the choice of functional form is important for the issue of identifying structural effects and for making inferences about the nature and the significance of such effects.

The choice of functional form is of particular importance in applications where the problem is to separate effects that stem from past behavior on current preferences from spurious effects related to correlation between current and past choices due to unobservables. This identification problem is crucial in a variety of contexts and it has been discussed most notably by Heckman (see Heckman, 1978, 1981a, 1981b, 1991, and the references therein). For example, in analyses of unemployment it is often noted that individuals who have experienced unemployment in the past are more likely to experience the event in the future than are individuals who have not experienced the event. As is well known, there may be two explanations for this empirical regularity. One explanation is that current and past choices are correlated due to unobservables that affect the preferences and which are serially correlated (pure taste persistence). In this case, past choices are proxies for unobserved variables that affect preference evaluations or unobservable opportunity sets, and consequently the aggregate transition rates will depend on past choices. The other explanation is that, as a result of choice experience, preferences and/or choice constraints change (structural state dependence). This fundamental identification problem cannot be solved without imposing theoretical restrictions in the model. This example illustrates the relevance of providing a theoretical rationale for the mathematical structure of econometric models of intertemporal discrete choice.

This paper addresses the problem of functional form in intertemporal discrete choice models. Specifically, we first propose a formal axiomatic characterization of intertemporal choice models under pure taste persistence. Since models with pure taste persistence represent a reference case it is important to characterize this case theoretically so as to provide a point of departure for specifying state dependence effects. Next, we discuss the extension to the case with state dependence. Our characterization of choice behavior under pure taste persistent preferences can be viewed as a stochastic formulation of rational behavior with exogenous preferences. The main theoretical assumption is in fact an intertemporal version of Luce's axiom "Independence from Irrelevant Alternatives". This assumption, together with some regularity conditions, imply that the choice process (as a process in time) becomes a Markov chain where the transition probabilities have a particular simple structure. Drawing on results obtained by Dagsvik (1983) and (1988) it follows that this Markov chain model is compatible with a random utility representation where the utilities associated with each alternative are independent extremal processes.

Assumptions that are analogous to the one described above were proposed by Dagsvik (1992) and (1995a).

The paper is organized as follows: In Section 2 the choice setting is formally described. In Section 3 the basic assumptions are introduced and discussed. Furthermore, we derive some important implications for the choice model. In Section 4 we demonstrate that our framework allows for an interpretation that is consistent with optimizing behavior in a life cycle context where the (chosen) expenditure path can be treated as if it were exogenous when analyzing intertemporal discrete choice and savings. In Section 5 we briefly discuss interpretations in the context of econometric specification of hazard functions and transition intensities. In Section 6 we analyze the case with time varying choice sets, and in Section 7 we consider the extension of the framework to allow for state dependence.

## **2. The choice setting**

The individual decision-maker (agent) is supposed to have preferences over a finite set of alternatives. The preferences are assumed random (even to the agent himself) in the sense that they vary from one moment in time to the next in a way that cannot fully be predicted by the agent. Although Georgescu-Roegen (1958), p. 158, argue that "we must begin with a theory of the individual not as a perfect choosing-instrument, but as a stochastic one", the notion of random preferences is not very common in economics where stochastic utilities usually are motivated by unobservables that are assumed perfectly foreseeable from the agent's viewpoint. In the psychological literature however, there is a long tradition dating back to Thurstone (1927) in which utilities are perceived as random to the agent. The reason for this is of course that individuals have been observed to behave inconsistently in laboratory choice experiments in the sense that a given agent makes different choices under identical experimental conditions (cf. Tversky, 1969). One explanation for this is that the agent is viewed as having difficulties with assessing the proper value (to him) of the choice alternatives (optimization error). The agent may therefore not have complete confidence in his judgment and feel that, in a different state of mind, he might have made a different choice (cf. Hogarth, 1982, Fischhoff et al., 1980, Tversky and Kahneman, 1983). Thus, while the agent (and the observing econometrician) are unable to predict future taste-shifters the taste-shifters realized in the past are known to the agent but unobserved to the econometrician.

Although the intra-individual randomness is stressed here the assumption of stochastic rationality introduced in Section 3 also allows the interpretation of utilities that are deterministic to the agent but random to the observer. However, the interpretation is less obvious in the case when the random utilities are deterministic to the agent.

Let  $S$  be the index set of  $m$  alternatives,  $a_1, a_2, \dots, a_m$ , and let  $\mathfrak{S}$  be the index set that corresponds to the collection of all non-empty subsets from  $S$ . To each alternative,  $a_j$ , there is associated a stochastic process,  $\{U_j(t), t \geq 0\}$ , where  $U_j(t)$  is the agent's (conditional indirect) utility of  $a_j$  given the information and choice history at time  $t$ . Moreover, each alternative,  $a_j$ , is characterized by an attribute vector,  $Z_j(t)$ , at time  $t$ . The vector  $Z_j(t)$  may also contain components that are interaction terms between attributes and agent-specific characteristics. The agent chooses  $a_j$  at age  $t$  if  $U_j(t)$  is the highest utility at  $t$ . Here age (time) is continuous. Let  $\{J(t)\} = \{J(t, B(t))\}$  denote the choice process, i.e.,

$$J(t) = j \text{ if } U_j(t) > \max_{k \neq j, k \in B(t)} U_k(t)$$

where  $\{B(t), t > 0, B(t) \in \mathfrak{S}\}$  denotes the choice set process. We define the choice set process to be increasing at time  $t$  if  $B(t) \setminus B(t-)$  is non-empty, and decreasing if  $B(t-) \setminus B(t)$  is non-empty. If  $B(t) = B(s)$  for all  $s$  and  $t$  the choice set process is constant. Let  $h(t) = \{J(s), s < t\}$  denote the choice history and define

$$U(t) = (U_1(t), U_2(t), \dots, U_m(t)),$$

and

$$Z(t) = (Z_1(t), Z_2(t), \dots, Z_m(t)).$$

We assume that the process  $Z \equiv \{Z(t), t > 0\}$  is exogenous. Furthermore, let  $t_r = (t_1, t_2, \dots, t_r)$  where  $t_1 < t_2 < \dots < t_r$ . We assume that  $\{U(t)\}$  is separable and continuous in probability.<sup>1</sup> Moreover, we assume that the cumulative distribution function (c.d.f.) of  $U(t)$  is absolutely continuous for any  $t \in \mathbb{R}_+$ . This implies that there are no ties, that is

$$P(U_i(t) = U_j(t)) = 0.$$

When the finite dimensional distributions have been specified it is in principle possible to derive joint choice probabilities for a sequence of choices. However, the class of intertemporal random utility models is quite large and it is thus of substantial interest to restrict this class on the basis of theoretical arguments. Also it seems to be rather difficult to find stochastic processes that are convenient candidates for utility processes in the sense that they imply tractable expressions for the choice probabilities in the intertemporal context.

### 3. Characterization of pure-taste-persistent preferences and choice probabilities

In the present section we propose behavioral assumptions that enable us to characterize preferences and the choice probabilities in the "reference state", where there are no effects from past experiences on future preferences nor on future choice opportunities. The extension to allow for state dependence is discussed in Section 7.

One way of introducing structural restrictions into the model is to apply probabilistic versions of the assumption of rational behavior. A famous example of this type of assumption is Luce Choice Axiom; "Independence from irrelevant alternatives", (IIA) (cf. Luce, 1959). A first attempt to extend IIA to the intertemporal setting was made by Dagsvik (1983).<sup>2</sup> Below we shall discuss the implications from another version of IIA, which will be introduced below. Let

$$F(y; t_r, Z, \gamma(h(t_r)), U(t_{r-1})) \equiv P(U(t_r) \leq y | U(t_i), Z, \forall i \leq r-1) \quad (3.1)$$

be the conditional marginal distribution of  $U(t_r)$  given  $Z$  and  $U(t_i)$ ,  $i=1,2,\dots,r-1$ . We assume that  $F$  is known apart from a parameter vector  $\gamma$  (say). The notation above signifies that the parameter vector  $\gamma$  that governs the conditional distribution of the utility processes, may change as a result of experience. It should be emphasized that in a strict mathematical sense, the left hand side of (3.1) is not immediately meaningful since when  $U(t_{r-1})$  is given,  $h(t_r)$  follows. But this regards precisely the core of the specification and identification problem, namely that without additional theory  $\gamma(\cdot)$  cannot be determined.

#### Definition 1

*By pure-taste-persistent preferences (PTPP) it is understood that  $\gamma$  in (3.1) is independent of  $h(t_r)$  for any  $t_r$ , i.e., there are no effects on the parameters of the current conditional c.d.f. of the agent's preferences from previous choices.*

Thus PTPP means that preferences are exogenous relative to the choice process.<sup>3</sup>

Let

$$P_j(t) \equiv P_j(t; B(t)) \equiv P\left(U_j(t) = \max_{k \in B(t)} U_k(t)\right) \quad (3.2)$$

and

$$Q_{ij}(s, t) \equiv Q_{ij}(s, t; B(t), h(s)) \equiv P\left(U_j(t) = \max_{k \in B(t)} U_k(t) | J(s) = i, h(s)\right) \quad (3.3)$$

for  $s < t$ . In general, the transition probabilities (given the choice history), will also depend on past choice sets and exogenous variables. For notational simplicity, this is suppressed in the definitions above.

We realize that  $Q_{ij}(s, t; B(t), h(s))$  is the conditional choice probability of choosing alternative  $a_j$  at time  $t$  given that alternative  $a_i$  was chosen at time  $s$  and given the choice history before  $s$ .

Provided the choice set does not change in a neighborhood of time  $t$ , we define the transition intensities,  $\{q_{ij}(t; B(t), h(t))\}$ , of  $\{J(t), t > 0\}$  by

$$q_{ij}(t; B(t), h(t)) \equiv \lim_{s \rightarrow t} \frac{Q_{ij}(s, t; B(t), h(s))}{t - s} \quad (3.4)$$

for  $i \neq j$ , and

$$q_{ii}(t; B(t), h(t)) \equiv \lim_{s \rightarrow t} \frac{(Q_{ii}(s, t; B(t), h(s)) - 1)}{t - s}. \quad (3.5)$$

Define finally

$$\pi_{ij}(t) \equiv \pi_{ij}(t; B(t), h(t)) \equiv \lim_{s \rightarrow t} \frac{Q_{ij}(s, t; B(t), h(s))}{\sum_{k \in B(t) \setminus \{i\}} Q_{ik}(s, t; B(t), h(s))} = \frac{q_{ij}(t; B(t), h(t))}{q_{ii}(t; B(t), h(t))}, \quad (3.6)$$

for  $i \neq j$ . The interpretation of (3.6) is as the transition probability of going to alternative  $a_j$  at time  $t$  given that alternative  $a_i$  is left and given the choice history prior to  $t$ .

**Assumption A1** (Stochastic rationality under PTPP)

Let  $B(s) = B(t-)$ , for  $s < t$ ,  $B(s) \in \mathfrak{S}$  and let  $B \in \mathfrak{S}$  be such that  $B \setminus B(t-) \neq \emptyset$ . Then, for  $j \in B \setminus B(t-)$ ,

$$Q_{ij}(t-, t; B, h(t-)) = P_j(t; B). \quad (3.7)$$

Assumption A1 states that provided the preferred alternative in  $B(t) = B$  at time  $t$  lies in  $B \setminus B(t-)$ , it does not depend (in a systematic way) on which of the alternative in  $B(t-)$  are preferred just before time  $t$ .

To facilitate interpretation and motivation we shall discuss A1 in the context of the following example, with  $B(t-) = \{1, 2\}$  and  $B = \{1, 2, 3\}$ . Then with  $i=1, 2$ , and  $j=3$ , (3.7) can in this example be expressed as

$$P(U_3(t) > \max(U_1(t), U_2(t)) | U_1(t-) = \max(U_1(t-), U_2(t-)), h(t)) = P(U_3(t) > \max(U_1(t), U_2(t))) \quad (3.8)$$

The left hand side of (3.8) expresses the probability that  $a_3$  is the preferred alternative given that  $a_2$  is the second preferred alternative and given the choice history. The right hand side of (3.8) is simply the unconditional probability that  $a_3$  is the preferred alternative. Thus (3.8), states that the event, " $a_3$  is the preferred alternative", is stochastically independent of the event, " $a_2$  is the preferred alternative among the remaining alternatives  $a_1$  and  $a_2$ ", and of the choice history. Another way of expressing this property goes as follows: The event, " $a_3$  is the preferred alternative from the set  $\{a_1, a_2, a_3\}$  at time  $t$ ", is stochastically independent of previous rankings of  $\{a_1, a_2\}$ . Thus we realize that A1 is a version of IIA that is analogous to versions studied by Luce (1959), and Block and Marschak (1960) concerning ranking theories (see Luce and Suppes, 1965, p.p. 353-354). It is therefore natural to interpret A1 as a characterization of stochastically rational agents under PTPP.

#### Assumption A2

*At each point in time the distribution of the random term,  $U(t)-EU(t)$ , does not depend on  $Z(t)$ .*

#### Assumption A3

*For any  $t > 0$ ,  $j \in S$ , and any real number  $x$  there exists a value of  $Z_j(t)$  such that  $EU_j(t)=x$ .*

#### Assumption A4

*Apart from a location shift the finite dimensional laws of the indirect utility process  $\{\max_k U_k(t), t \geq 0\}$  are the same as the finite dimensional laws of  $\{U_1(t), t \geq 0\}$ .*

Assumption A2 states that at each moment in time the random term of the utility function is independent of the structural term. Assumption A3 states that the structural term of the utility function can vary over the whole real line when attributes vary freely.

Recall that the max-stable processes have the property that maximum of independent max-stable processes is max-stable (see de Haan, 1984). The finite dimensional distributions of a multivariate max-stable process belong to the class of multivariate extreme value distributions.

Dagsvik (1995b) has demonstrated that there is no loss of generality in assuming A4 since, in the absence of structural state dependence effects, any intertemporal random utility model can, under suitable regularity conditions, be approximated arbitrarily closely by choice probabilities generated from max-stable utilities.

### Theorem 1

Assume that A2 and A3 hold. Then for any  $B \in \mathfrak{S}$ , Assumption A1 implies that

$$P_j(t) \equiv P_j(t; B) = \frac{e^{v_j(t)}}{\sum_{k \in B} e^{v_k(t)}} \quad (3.9)$$

where  $v_j(t) = \alpha EU_j(t)$  and  $\alpha > 0$  is an arbitrary constant.

### Proof:

Recall that  $\{U(t), t > 0\}$  is continuous in probability. Note that with  $B(t-) = B \setminus \{j\}$ , A1 implies that

$$P\left(U_j(t) = \max_{k \in B} U_k(t), U_i(t-) = \max_{k \in B \setminus \{j\}} U_k(t-)\right) = P\left(U_j(t) = \max_{k \in B} U_k(t)\right) P\left(U_i(t) = \max_{k \in B \setminus \{j\}} U_k(t)\right). \quad (3.10)$$

By Theorem 50, p. 354, in Luce and Suppes (1965), (3.10) implies that  $P_j(B)$  is a strict utility model as expressed in (3.9). Finally, Strauss (1979), p.p. 42-43, has demonstrated that the parameters  $\{v_j(t)\}$  of the choice model are related to the utility function by  $v_j(t) = \alpha EU_j(t)$ , apart from an additive constant.

Q.E.D.

### Remark

Without loss of generality we shall in the following put  $\alpha=1$ .

Let us now proceed by investigating the intertemporal structure of the random utilities that follows from A1.<sup>4</sup>

Above we postulated the existence of random utility processes such that A1 to A3 hold. It remains, however, to demonstrate that such processes really exist. In the one-period case McFadden (1973), Yellott (1977) and Strauss (1979) have, under different sets of conditions, demonstrated the equivalence between IIA and extreme value distributed utilities in a random utility model with independent utilities. We state a version of this result in the next theorem.

### Theorem 2

Assume that A1 to A4 hold. Then the utility processes,  $\{U_j(t), t \geq 0\}$ ,  $j=1,2,\dots,m$ , are independent at each point in time and have type III extreme value distributed marginals.<sup>5</sup>

### Proof:

It follows from A4 that the joint distribution of  $(U_1(t), U_2(t), \dots, U_m(t))$  belongs to the class of c.d.f. considered in Theorem 6 of Strauss (1979) with  $\varphi(x) = e^{-x}$ . The result now follows from Strauss, Theorem 6.<sup>6</sup>

Q.E.D.

### Assumption A5

The utility processes  $\{U_j(t), t \geq 0\}$ ,  $j = 1, 2, \dots, m$ , are stochastically independent.

Note that two stochastic processes  $\{U_i(t), t \geq 0\}$  and  $\{U_j(t), t \geq 0\}$  may be stochastically dependent even if  $U_i(t)$  and  $U_j(t)$  are stochastically independent at each point in time. For example,  $U_i(t)$  and  $U_j(s)$  may be interdependent for  $s \neq t$  even if  $U_i(t)$  and  $U_j(t)$  are independent. However, it may be plausible in many applications to assume that the correlation between  $U_i(t)$  and  $U_j(s)$  is less than the correlation between  $U_i(t)$  and  $U_j(t)$ , which implies that the utility processes are independent when the utilities at each point in time are independent.

### Theorem 3

Assume A1 to A5. Then the utilities are extremal processes with type III extreme value marginal distribution.

The proof of Theorem 3 is given in the appendix.

The class of extremal processes was introduced in statistics by Dwass (1964, 1966, 1974) and Tiago de Oliveira (1968, 1973). An extension to inhomogeneous extremal processes has been made by Weissman (1975). Let us briefly review the properties of what is called extremal processes with type III extreme value marginal distributions, denoted by  $\{Y(t), t \geq 0\}$  (say). This process has the same law as  $\{\tilde{Y}(t), t \geq 0\}$  defined by

$$\tilde{Y}(t) = \max(\tilde{Y}(s), \tilde{W}(s, t)), \quad (3.11)$$

$s < t$ , where  $\tilde{Y}(0) = -\infty$  and where  $\tilde{W}(s, t)$  is independent of  $\tilde{Y}(s)$  and has type III extreme value c.d.f. for  $s < t$ . Moreover,  $\tilde{W}(s, t)$  and  $\tilde{W}(s', t')$  are independent when  $(s, t) \cap (s', t') = \emptyset$ . For a more detailed description of extremal processes we refer to Resnick (1987). It follows from (3.11) that an extremal process is a Markov process. Moreover, it is a pure jump Markov process.

From (3.11) it also follows that the extremal process is non-decreasing with probability one. For the sake of interpretation it may be more intuitive to apply a "detrended" version of the extremal process as a candidate for a utility representation. To this end we shall assume that

$\{U_j(t) + \beta(t), t \geq 0\}$  is an extremal process, where  $\{U_j(t), t \geq 0\}$  is the utility process associated with alternative  $j$  and  $\beta(t)$  is a suitable nondecreasing function. Analogous to (3.11) the detrended utility process can be represented as

$$U_j(t) = \max(U_j(s) + \beta(s) - \beta(t), W_j(s, t)) \quad (3.12)$$

for  $s < t$ , where  $U_j(0) = -\infty$  and where  $\{W_j(s, t)\}$  has similar properties as  $\{\tilde{W}(s, t)\}$ , that is  $W_j(s, t)$  and  $W_j(s', t')$  are independent when  $(s, t) \cap (s', t') = \emptyset$ , with type III extreme value c.d.f. One can express the mean of  $W_j(s, t)$  as

$$\exp(E W_j(s, t)) = \int_s^t \exp(w_j(\tau)) d\tau \quad (3.13)$$

where  $w_j(\tau)$  is a suitable function. From (3.11), (3.12) and (3.13) we obtain readily that

$$\exp(E U_j(t) + \beta(t)) - \exp(E U_j(t - \Delta t) + \beta(t - \Delta t)) = \Delta t \exp(w_j(t) + \beta(t)) + o(\Delta t)$$

which implies that

$$\exp(E U_j(t) + \beta(t)) = \int_0^t (\exp(w_j(\tau) + \beta(\tau))) d\tau,$$

or equivalently

$$\exp(E U_j(t)) = \int_0^t (\exp(w_j(\tau) + \beta(\tau) - \beta(t))) d\tau. \quad (3.14)$$

When  $w_j(t)$  is independent of time and  $\beta(t) = t\theta$ , where  $\theta > 0$  is a constant, (3.14) reduces to

$$\exp(E U_j(t)) = e^{w_j} (1 - e^{-\theta t}) / \theta. \quad (3.15)$$

Thus for large  $t$ , a constant mean utility level corresponds to constant  $\{w_j(t)\}$  when  $\beta(t)$  is linear in  $t$ . Also from (3.14) we realize that when  $\beta(t) = t\theta$ ,  $\theta$  is analogous to a rate of preference parameter because by (3.14), the mean utility at time  $t$  can be expressed as a weighted integral of past "instantaneous" mean utilities. Specifically, the contribution from the period  $s$ -specific systematic utility component to the current mean utility is evaluated by multiplying  $\exp(w_j(s))ds$  by the "depreciation" factor  $\exp(-(t-s)\theta)$ .<sup>7</sup> This depreciation factor accounts for the loss of memory and/or decrease in taste persistence as the time lag increase.

As demonstrated by Resnick and Roy (1990) we can express a particular version of the autocorrelation function of the utility process as

$$\text{corr}\left(\exp(-U_j(s)), \exp(-U_j(t))\right) = \frac{\exp(EU_j(s))}{\exp(EU_j(t))} \exp(\beta(s) - \beta(t)). \quad (3.16)$$

for  $s < t$ . To clarify the interpretation, consider the case with  $\beta(t) = t\theta$ , where  $\theta > 0$  is a constant. Then (3.16) reduces to

$$\text{corr}\left(\exp(-EU_j(s)), \exp(-EU_j(t))\right) = \frac{1 - e^{-\theta s}}{1 - e^{-\theta t}} \cdot e^{-(t-s)\theta}. \quad (3.17)$$

Thus when  $s$  and  $t$  are large the mean utility in this case equals  $w_j$ , (apart from an additive constant) and the auto-correlation function becomes exponential.

#### Definition 2

We define a modified extremal process as a stochastic process which satisfies (3.12) with  $U_j(0) = -\infty$ , with a nondecreasing function  $\beta(\cdot)$ , and with type III extreme value distributed marginals.

#### Theorem 4

Assume that the random utilities are independent modified extremal processes. Assume furthermore that the choice set process is constant over time. Then (3.9) holds and  $\{J(t), t > 0\}$  is a Markov chain. Furthermore, the transition probabilities are given by

$$Q_{ij}(s, t) = P_j(t) - \zeta(s, t) P_j(s) \quad (3.18)$$

for  $i \neq j, s < t$  and

$$Q_{ii}(s, t) = P_i(t) + \zeta(s, t) (1 - P_i(s)) \quad (3.19)$$

where  $v_j(t) = EU_j(t)$ ,  $B \in \mathfrak{S}$ ,  $P_j(t)$  is given by (3.9), and

$$\zeta(s, t) = \frac{\sum_{k \in B} e^{v_k(s) + \beta(s) - \beta(t)}}{\sum_{k \in B} e^{v_k(t)}}. \quad (3.20)$$

A proof of Theorem 4 is given in Dagsvik (1988).

Theorem 4 shows that A1 to A4 imply strong testable restrictions since — in addition to the Markov property — the transition probabilities are independent of  $i$  when  $i \neq j$ .

#### Corollary 1

Assume that  $\{\beta(t) + v_j(t)\}$  are differentiable with respect to  $t$ . Under the assumptions of Theorem 4 the transition intensities of the Markov chain  $\{J(t), t > 0\}$  exists and are given by

$$q_{ij}(t) = P_j(t) (v'_j(t) + \beta'(t)) \quad (3.21)$$

for  $i \neq j$ , and

$$q_{ii}(t) = - \sum_{k \in B \setminus \{i\}} q_{ik}(t) \quad (3.22)$$

where  $P_j(t)$  is given by (3.9). The probability of going to state (alternative)  $a_j$  given that state (alternative)  $a_i$  is left, equals

$$\pi_{ij}(t) = \frac{e^{v_j(t)} (v'_j(t) + \beta'(t))}{\sum_{k \in B \setminus \{i\}} e^{v_k(t)} (v'_k(t) + \beta'(t))} \quad (3.23)$$

where  $v'_j(t) + \beta'(t)$  denotes the derivative of  $v_j(t) + \beta(t)$ .

The results of Corollary 1 follow directly from (3.9), (3.18), (3.19) and (3.20).

#### Corollary 2

Under the assumptions of Theorem 4 the indirect utility,  $\max_{k \in B} U_k(t)$ , is independent of  $\{J(\tau), \tau \leq t\}$  for any  $B \in \mathfrak{S}$ .

A proof of this result has been given by Resnick and Roy (1990).

Dagsvik (1988) and Resnick and Roy (1990) extend the result of Theorem 4 to the case where  $\{U(t), t \geq 0\}$  is a multivariate extremal process. Dagsvik considers the case where  $U(t)$  — at each  $t$  — has a type III multivariate extreme value distribution that is absolutely continuous. The resulting (marginal) choice probabilities at a given point in time in this case become generalized extreme value probabilities. Resnick and Roy (1990) allow  $U(t)$  to have a multivariate c.d.f. that is not necessarily absolutely continuous.

Recall that by (3.12) the utility processes are Markov processes. However, utility processes with the Markov property do not usually imply that the corresponding choice process  $\{J(t)\}$  is Markovian. For example, Gaussian utility processes with the Markovian property do not imply that the choice process is Markovian. In fact, there exist no Gaussian utility processes in continuous time that can generate Markovian choice models.

Similarly, to (3.16) we have that since the indirect utility,  $\max_{k \in B} U_k(t)$ , is a modified extremal process with

$$\exp\left(E\left(\max_{k \in B} U_k(t)\right)\right) = \sum_{k \in B} \exp(EU_k(t))$$

it follows that

$$\text{corr}\left(\exp\left(-\max_{k \in B} U_k(s)\right), \exp\left(-\max_{k \in B} U_k(t)\right)\right) = \zeta(s, t). \quad (3.24)$$

In other words,  $\zeta(s, t)$  represents the autocorrelation function of the indirect utility,  $\max_{k \in B} U_k(t)$ .<sup>8</sup>

#### 4. Life cycle consistent choice behavior

In this section we consider the following setting: The agent must make a choice between  $m$  different alternatives (states) in each period (time is discrete). Let  $\alpha_{jt}$  be a period specific cost or income variable associated with alternative  $j$  and  $c_t$  the (composite) consumption in period  $t$ . There are no transaction costs and the preferences are assumed to be exogenous (i.e. there is no state dependence). The extension to the case with state dependence will be discussed in Section 7. Let  $y_t$  denote total expenditure in period  $t$  and let  $r_t$  be the interest rate in period  $t$ . The price index is equal to one. Furthermore, let  $\omega_t$  be the income in period  $t$  and  $Y_t$  the wealth at the end of period  $t$ .

The budget constraints in period  $t$  are given by

$$y_t = Y_t + \omega_t - \frac{Y_{t+1}}{1+r_{t+1}} - T(\omega_t, Y_t) \quad (4.1)$$

and

$$y_t = \sum_i 1_{\{J(t)=i\}} \alpha_{it} + c_t \quad (4.2)$$

where  $T(\cdot)$  is the tax function. Let  $U_j(t, c_t)$  denote the instantaneous utility as of period  $t$  given  $c_t$  and given that  $J(t)=j$ . Future incomes, interest rates and costs are uncertain. Let  $V_j(t, Y_t)$  denote the value function as of period  $t$  given that  $J(t)=j$ . Under the assumption of additive intertemporal separability, the Bellman equation that corresponds to the dynamic optimization problem described above is given by

$$V_j(t, Y_t) = U_j(t, y_t - \alpha_{jt}) + \rho E_t \left( \max_k V_k(t+1, Y_{t+1}) \mid J(t)=j \right) \quad (4.3)$$

where  $\rho \leq 1$ , is the time-preference discounting factor and  $E_t$  denotes the subjective expectation operator given the agent's information at time  $t$ . Since we assume that preferences are exogenous it follows that

$$E_t \left( \max_k V_k(t+1, Y_{t+1}) \mid J(t)=j \right) = E_t \left( \max_k V_k(t+1, Y_{t+1}) \right). \quad (4.4)$$

This means that  $J(t)$  is determined by maximizing  $U_j(t, Y_t - \alpha_{jt})$ . Let

$$U(t, y) \equiv U(t, y, \alpha_t) \equiv \max_k U_k(t, y - \alpha_{kt})$$

and

$$V(t, Y) \equiv \max_k V_k(t, Y).$$

By (4.3) and (4.4) it follows that  $y_t$  is determined as

$$y_t = \arg \max_y \left( U(t, y) + \rho E_t V(t+1, (1+r_{t+1})(Y_t + \omega_t - T(\omega_t, Y_t) - y)) \right) \quad (4.5)$$

and

$$V(t, Y_t) = U(t, y_t) + \rho E_t V(t+1, Y_{t+1}). \quad (4.6)$$

From (4.1) and (4.5) it follows by recursion that  $V(t, Y_t)$  depends on  $Y_t$  and the subjective expectation of respective future discounted instantaneous indirect utilities evaluated at the optimum expenditure path where it is understood that future expenditures are evaluated conditional on the information that becomes available at the times the respective future decisions are being made. Thus, the decision problem can be viewed as a two stage process in which the agent determines the

expenditure path in the first stage and the optimal choice of state in each period is determined conditional on the expenditure path. So far, however, we have just reviewed a version of the wellknown two stage decomposition (two stage budgeting) of the intertemporal decision problem. As is also wellknown, the application of two stage budgeting in empirical analyses may be difficult because the optimal expenditure path may in general be correlated with the random terms of  $\{U_j(t, y), j = 1, 2, \dots, t = 1, 2, \dots\}$ . However, due to the properties of the extremal processes we shall see in a moment that this will not be the case here.

Let  $\{x_t\}$  be any given sequence of real numbers. Suppose  $\{U_j(t, x_t), t > 0\}, j = 1, 2, \dots, m$ , are independent modified extremal processes. Consistent with the notation in Section 2, let

$$J(t, y) = j \Leftrightarrow U_j(t, y - \alpha_{jt}) = U(t, y). \quad (4.7)$$

By Corollary 2 it follows that  $\{J(\tau, x_\tau), \tau \leq t\}$  is independent of  $\{U(\tau, x_\tau), \tau \geq t\}$  for any  $\{x_\tau\}$ . Note furthermore that when  $Y_t$  is given then  $\{y_\tau, \tau \geq t\}$  is — by (4.5) and (4.6) — determined from  $\{U(\tau, \cdot), \tau \geq t\}$  and by the subjective expectation operator. Suppose also that the agent's subjectively perceived law of  $\{U(\tau, x_\tau), \tau \geq t\}$  is similar to the corresponding objective law in that  $\{U(\tau, x_\tau), \tau \geq t\}$  and  $\{J(\tau, x_\tau), \tau \leq t\}$  are independent. This would of course be true if the agent's subjective law equals the objective one (rational expectation) — or if the agent knows future taste-shifters with perfect certainty. Since  $\{y_\tau, \tau \geq t\}$  is determined by  $Y_t$ , and by the distribution of  $\{U(\tau, \cdot), \tau \geq t\}$ , future interest rates and incomes, it follows that  $\{y_\tau, \tau \geq t\}$  must be independent of  $\{J(\tau, y_\tau), \tau \leq t\}$ . But this means that, under the assumptions of Theorem 4, we can analyze the choice process  $\{J(t, y_t), t > 0\}$  conditional on  $\{y_t\}$  as if  $\{y_t\}$  were exogenous.

## 5. Some implications for econometric specifications of transition intensities under pure taste persistence

The results in Theorem 4 and Corollary 1 are useful for justifying the choice of functional form of the likelihood function of observations on  $\{J(\tau), Z(\tau), \tau \leq t\}$  for a particular agent under PTPP. The first step in specifying an empirical model is to specify the structural parts of the model.

Recall that according to the representation of the modified extremal process, (3.13) yields  $\exp(EW_j(t - dt, t)) = \exp(w_j(t))dt$  which means that  $w_j(t)$  is equivalent to the mean of the increment  $W_j(t - dt, t)$  at time  $t$ . Moreover, we noticed above that due to (3.13) it is possible to express the

structural term of the current utility as a "depreciated" sum of the structural parts of the past increments. This allows us to interpret  $w_j(t)$ , or equivalently  $\exp(w_j(t))$ , as the representative instantaneous utility of alternative  $a_j$  at time  $t$ . In empirical applications one would typically specify  $w_j(t)$  as

$$w_j(t) = w(Z_j(t)) \quad (5.1)$$

where  $w(\cdot)$  is a suitably chosen functional form that is known apart from an unknown vector of parameters. If we assume  $\beta(t) = \theta t$ ,  $\theta > 0$  and insert (3.13) for  $v_j(t)$  in Theorem 4 we obtain the next result:

### Corollary 3

*Assume that the utilities are independent modified extremal processes with  $\beta(t) = t\theta$ ,  $\theta \geq 0$ , and the choice sets process is constant over time. Then, under the Assumptions of Theorem 4*

$$P_j(t) = \frac{\int_0^t e^{w_j(\tau) - (t-\tau)\theta} d\tau}{\sum_{k \in B} \int_0^t e^{w_k(\tau) - (t-\tau)\theta} d\tau}, \quad (5.2)$$

$$Q_{ij}(s, t) = \frac{\int_s^t e^{w_j(\tau) - (t-\tau)\theta} d\tau}{\sum_{k \in B} \int_0^t e^{w_k(\tau) - (t-\tau)\theta} d\tau}, \quad (5.3)$$

$$q_{ij}(t) = \frac{e^{w_j(t)}}{\sum_{k \in B} \int_0^t e^{w_k(\tau) - (t-\tau)\theta} d\tau}, \quad (5.4)$$

$$\pi_{ij}(t) = \frac{e^{w_j(t)}}{\sum_{k \in B \setminus \{i\}} e^{w_k(t)}} \quad (5.5)$$

for  $i \neq j$ , and

$$\zeta(s, t) = \frac{\sum_{k \in B} \int_0^s e^{w_k(\tau) - (s-\tau)\theta} d\tau}{\sum_{k \in B} \int_0^t e^{w_k(\tau) - (t-\tau)\theta} d\tau} e^{-(t-s)\theta}. \quad (5.6)$$

Evidently,  $q_{ii}(t)$  and  $Q_{ii}(s, t)$  are found from the adding-up conditions.

Let us next consider the particular case where  $\{w_j(t)\}$ ,  $j = 1, 2, \dots, m$ , are constant over time i.e.,  $w_j(t) = w_j$ . Then (5.4) and (5.6) reduce to

$$q_{ij}(t) = \frac{\theta e^{w_j}}{(1 - e^{-t\theta}) \sum_{k \in B} e^{w_k}} \equiv \frac{\theta P_j}{1 - e^{-t\theta}} \quad (5.7)$$

and

$$\zeta(s, t) = \frac{(1 - e^{-s\theta}) e^{-(t-s)\theta}}{1 - e^{-t\theta}}. \quad (5.8)$$

From (3.24) it follows that the degree of taste persistence in the indirect utility can be measured by  $\theta$ . Specifically, when  $\theta$  is large there is little taste persistence (provided  $s$  and  $t$  are large) while when  $\theta$  is close to zero tastes are strongly correlated over time. Moreover, (5.7) shows that the transition intensities are stationary when  $t$  is large. However, when  $t$  is small then the transition intensities given by (5.7) depend on time. This is due to the fact that in the beginning of a choice process the length of the choice history (age) will influence the strength of the taste persistence effect.

Observe that the structure of (5.7) can be viewed as a special case of the model in Olsen et al. (1986). However, as the utilities in their model are not serially correlated,  $\theta$ , in their model seems at first glance to yield a different interpretation. In their framework the utilities are viewed as independent draws that occur according to a Poisson process with intensity  $\theta$ . But this means in fact that also in their setting  $\theta$  allows the interpretation as a measure of taste persistence because when  $\theta$  is small the random draws occur rarely and therefore preferences are rather stable over time. In contrast when  $\theta$  is large preferences are likely to change frequently.

Let us finally compare the structure of the hazard rate with the proportional hazard rate framework which has been used extensively in empirical analyses, cf. Heckman and Singer (1985). In the case with time invariant explanatory variables and  $w_j(\tau) = w_j + \kappa(\tau)$ , where  $\kappa(\tau)$  is a function of time that is independent of the explanatory variables, we obtain the hazard rate from (3.5) and (5.4):

$$-q_{ii}(t) = \lambda_0(t)(1 - P_i) \quad (5.9)$$

where  $\lambda_0$  is given by

$$\lambda_0(t) \equiv \frac{e^{\theta t + \kappa(t)}}{\int_0^t e^{\theta \tau + \kappa(\tau)} d\tau}. \quad (5.10)$$

The term  $\lambda_0(t)$  corresponds to the so-called baseline hazard in the statistical literature. This shows that the proportional hazard rate assumption is consistent with the present framework. When  $\kappa(t)=0$ , and  $t$  is large, (5.10) reduces to  $\lambda_0(t)=\theta$ . By means of (5.6),  $\lambda_0(t)$  can be given a particular interpretation.

Under the present specification it follows from (5.6) and (5.10) that

$$\lambda_0(t) = \left. \frac{\partial \zeta(s, t)}{\partial s} \right|_{s=t}. \quad (5.11)$$

Consequently,  $\lambda_0(t)$  can be interpreted as a measure of the instantaneous change in the preferences in the neighborhood of  $t$ , due to random variation of the taste-shifters.

In the multistate case some authors (see for example Andersen et al. (1991)) have specified transition intensities  $\{\tilde{q}_{ij}(t)\}$  on the form

$$\tilde{q}_{ij}(t) = \lambda_{ij}(t) \exp\left(f(Z_i, Z_j; b)\right) \quad (5.12)$$

where  $f(\cdot)$  is some specified function,  $Z_i$  is an individual specific time invariant vector of covariates that characterize alternative  $a_i$ , and  $b$  is a vector of parameters. Let us now compare the structure (5.12) with (5.7) and (5.9). We realize that (5.7) and (5.9) are essentially different from (5.12) in that (5.7) and (5.9) depend on all the covariates in a particular way while (5.12) only depends on the covariates related to alternatives  $a_i$  and  $a_j$ . Therefore, the standard proportional hazard specification (5.12), which is often applied in duration analysis, is inconsistent with a random utility formulation when the number of states is larger than two.

## 6. Allowing for time varying choice sets

In many applications it is of interest to allow for time varying choice sets. For example, in the analysis of labor market dynamics, workers' market opportunities may depend on experience and possibly on previous unemployment spells. When modeling fertility histories one must take into account that a woman with — say — one child at most have the choice between getting an additional child or have no additional children.

It shall always be understood in the following that the choice sets can change at most a finite number of times.

When the choice sets vary over time the corresponding choice model will in general not be Markovian. We shall in this section discuss the choice process in this case. For the sake of interpretation, but with no loss of generality, we shall assume below that the trend function  $\beta(t)$  is linear.

**Theorem 5**

Assume that the utilities are modified extremal processes with  $\beta(t)=\theta t, \theta \geq 0$ . If the choice set process does not change at time  $t$ , then

$$q_{ij}(t, h(t)) = \frac{e^{w_j(t)}}{\sum_{k \in B(t)} \int_0^t e^{w_k(\tau) - (t-\tau)\theta} d\tau} \quad (6.1)$$

for  $i \neq j, i, j \in B(t) \in \mathfrak{S}$ .

If we compare (6.1) with (5.4) we realize that under the conditions of Theorem 5 the transition intensities (6.1) have the same structure as in the case with constant choice sets over time.

The result of Theorem 5 follows immediately from Lemma 4, which is stated and proved in the appendix.

**Theorem 6**

Assume that the utilities are modified extremal processes with  $\beta(t)=\theta t, \theta \geq 0$ . Suppose, moreover, that the choice set process increases at time  $t$ . Let  $r(t)$  be the last time before  $t$  for which  $J(r(t)) \in B(t) \setminus B(t-)$ . Then, if  $J(r(t)) \neq j$ ,

$$Q_{ij}(t-, t, h(t)) = \frac{\int_0^{r(t)} e^{w_j(\tau) - (t-\tau)\theta} d\tau}{\sum_{k \in B(t)} \int_0^t e^{w_k(\tau) - (t-\tau)\theta} d\tau} \quad (6.2)$$

for  $i \in B(t-), j \in B(t) \setminus B(t-)$ . If  $J(r(t)) = j$ , then

$$Q_{ij}(t-, t, h(t)) = 1 - \frac{\sum_{k \in B(t) \setminus \{j\}} \int_0^{r(t)} e^{w_k(\tau) - (t-\tau)\theta} d\tau}{\sum_{k \in B(t)} \int_0^t e^{w_k(\tau) - (t-\tau)\theta} d\tau} \quad (6.3)$$

for  $j \in B(t) \setminus B(t-)$ . For  $i \neq j$ ,  $i, j \in B(t-)$ ,  $Q_{ij}(t-, t, h(t)) = 0$ , and

$$Q_{ii}(t-, t, h(t)) = 1 - \sum_{k \in B(t) \setminus B(t-)} Q_{ik}(t-, t, h(t)) \quad (6.4)$$

for  $i \in B(t-)$ .

The proof of Theorem 6 is given in the appendix.

For the sake of interpretation it is interesting to consider the special case when  $w_j(t)$  is constant over time for all  $j$ .

Corollary 4

Assume that  $w_j(t)$  is constant over time for all  $j$ . Then, under the assumptions of Theorem 6

$$Q_{ij}(t-, t, h(t)) = \frac{e^{w_j}}{\sum_{k \in B(t)} e^{w_k}} (1 - \exp(-(t - r(t))\theta)) \quad (6.5)$$

for  $i \in B(t-)$ ,  $j \in B(t) \setminus B(t-)$ , provided  $J(r(t)) \neq j$ . If  $J(r(t)) = j$ , then

$$Q_{ij}(t-, t, h(t)) = \exp(-(t - r(t))\theta) + \frac{e^{w_j}}{\sum_{k \in B(t)} e^{w_k}} (1 - \exp(-(t - r(t))\theta)) \quad (6.6)$$

for  $j \in B(t) \setminus B(t-)$ ,  $i \in B(t-)$ .

Remark

If  $J(s) \notin B(t) \setminus B(t-)$  for all  $s < t$  then  $r(t) = 0$ .

From (6.5) and (6.6) we see that the probability of moving from  $i$  to  $j$  is greater when  $J(r(t)) = j$  than when  $J(r(t)) \neq j$ . It is easily verified that this is true also in the general case stated in Theorem 6. The reason for this is that when  $J(r(t)) = j$ , this means a higher preference for alternative  $a_j$  than for any other feasible alternative at time  $r(t)$ . Since the autocorrelations of the utility processes are positive this means that the preference for alternative  $a_j$  at time  $t$  is likely to be higher when  $J(r(t)) = j$  than when  $J(r(t)) \neq j$ .

**Theorem 7**

Assume that the utilities are modified extremal processes with  $\beta(t)=\theta t, \theta \geq 0$ . Assume moreover that the choice set process is nondecreasing. Then the choice process  $\{J(t), t > 0\}$  is a Markov chain. The state and transition probabilities are given by

$$P_j(t) = \frac{\int_0^t e^{w_j(\tau) - (t-\tau)\theta} d\tau}{\sum_{k \in B(t)} \int_0^t e^{w_k(\tau) - (t-\tau)\theta} d\tau} \quad (6.7)$$

and

$$Q_{ij}(s, t) = P_j(t) - \zeta(s, t, B(s), B(t)) P_j(s) \delta_i(B(t)) \quad (6.8)$$

for  $i \neq j, i \in B(t), j \in B(t), B(s), B(t) \in \mathfrak{S}$ , where

$$\zeta(s, t, B(s), B(t)) = \frac{\sum_{k \in B(s)} \int_0^s e^{w_k(\tau) - (t-\tau)\theta} d\tau}{\sum_{k \in B(t)} \int_0^t e^{w_k(\tau) - (t-\tau)\theta} d\tau} \cdot e^{-(t-s)\theta},$$

and  $\delta_i(B(t)) = 1$  if  $i \in B(t)$ , and zero otherwise.

**Proof:**

The results of Theorem 7 follow from Resnick and Roy (1990) and from (3.19).

The next result concerns the case where the current choice set decreases.

**Theorem 8**

Assume that the utilities are modified extremal processes with  $\beta(t)=\theta t, \theta \geq 0$ , and suppose that the choice set process decreases at time  $t$ . Let  $s(t)$  be the last time before  $t$  for which  $J(s(t)) \in B(t)$ . Then

$$q_{ij}(t, h(t)) = \frac{e^{w_j(t)}}{\sum_{k \in B(t)} \int_0^t e^{w_k(\tau) - (t-\tau)\theta} d\tau} \quad (6.9)$$

for  $i \neq j, i, j \in B(t) \subset B(t-), B(t-), B(t) \in \mathfrak{S}$  and

$$q_{ii}(t, h(t)) = - \sum_{k \in B(t) \setminus \{i\}} q_{ik}(t, h(t)). \quad (6.10)$$

If  $J(s(t)) \neq j$ , then

$$Q_{ij}(t-, t, h(t)) = \frac{\int_0^{s(t)} e^{w_j(\tau) - (t-\tau)\theta} d\tau}{\sum_{k \in B(t)} \int_0^t e^{w_k(\tau) - (t-\tau)\theta} d\tau} \quad (6.11)$$

for  $i \in B(t-) \setminus B(t)$ ,  $j \in B(t)$ . If  $J(s(t)) = j$ , then

$$Q_{ij}(t-, t, h(t)) = 1 - \frac{\sum_{k \in B(t) \setminus \{i\}} \int_{s(t)}^t e^{w_k(\tau) - (t-\tau)\theta} d\tau}{\sum_{k \in B(t)} \int_0^t e^{w_k(\tau) - (t-\tau)\theta} d\tau} \quad (6.12)$$

for  $i \in B(t-) \setminus B(t)$ ,  $j \in B(t)$ .

A proof of Theorem 8 is given in the appendix.

As above, we realize that a transition probability from  $i$  to  $j$  is higher when  $J(s(t)) = j$  than when  $J(s(t)) \neq j$ .

The next result is immediate.

#### Corollary 5

Assume that  $w_j(t)$  is constant over time for all  $j$ . Then, under the assumptions of Theorem 8 we get that

$$Q_{ij}(t-, t, h(t)) = \frac{e^{w_j}}{\sum_{k \in B(t)} e^{w_k}} (1 - \exp(-(t - s(t))\theta)) \quad (6.13)$$

for  $i \in B(t-) \setminus B(t)$ ,  $j \in B(t)$ , provided  $J(s(t)) \neq j$ . If  $J(s(t)) = j$ , then

$$Q_{ij}(t-, t, h(t)) = \exp(-(t - s(t))\theta) + \frac{e^{w_j}}{\sum_{k \in B(t)} e^{w_k}} (1 - \exp(-(t - s(t))\theta)) \quad (6.14)$$

for  $i \in B(t-) \setminus B(t)$ ,  $j \in B(t)$ .

## 7. Extending the model to allow for state dependence

So far we have only discussed the functional form of the choice probabilities of  $\{J(t)\}$  under PTPP. The question now arises how the particular functional form that follows from PTPP should be modified in the presence of state dependence.

Notice first that when the utility processes are altered by the choice history a simultaneous equation bias problem arises. This is so because the structural terms of the utility processes become dependent on past choices, and consequently they will depend on past realizations of the utility processes.

For simplicity we shall consider the discrete time case. Accordingly, we will assume that the utility processes are independent modified experience-dependent extremal processes defined by

$$U_j(t) = \max(U_j(t-1) + \beta(t-1) - \beta(t), W_j(t, h(t))) \quad (7.1)$$

where

$$P(W_j(t, h(t)) \leq w | U(t-1)) = \exp(-\exp(g_j(t, h(t)) - w)) \quad (7.2)$$

and where  $g_j(t, h(t))$  is a parametric function of the attributes of alternative  $j$  and past choice experience. Define  $v_j(t, h(t))$  recursively by

$$\exp(v_j(t, h(t)) + \beta(t)) = \exp(v_j(t-1, h(t-1)) + \beta(t-1)) + \exp(g_j(t, h(t))). \quad (7.3)$$

Note that when  $g_j$  does not depend on  $h(t)$  then  $v_j(t, h(t))$  reduces to  $v_j(t) = EU_j(t)$ .

The following result extends Theorem 7 to the case with state dependence.

### Theorem 9

*Assume that the choice set process is non-decreasing and  $\{U_j(t), t \geq 0\}$ ,  $j = 1, 2, \dots, m$ , are independent and experience-dependent utility processes defined by (7.1) and (7.2). Then the (one step) transition probabilities, conditional on the choice history, are given by*

$$\begin{aligned} Q_{ij}(t-1, t, h(t)) &= R_j(t, h(t)) \left[ 1 - \exp(v_j(t-1, h(t-1)) - v_j(t, h(t)) + \beta(t-1) - \beta(t)) \right] \\ &\equiv R_j(t, h(t)) - \zeta(t-1, t, h(t)) R_j(t-1, h(t-1)) \delta_i(B(t)) \end{aligned} \quad (7.4)$$

for  $i \neq j$ ,  $j \in B(t)$ ,  $i \in B(t-1)$ , and

$$Q_{ii}(t-1, t, h(t)) = R_i(t, h(t)) + \zeta(t-1, t, h(t)) (1 - R_i(t-1, h(t-1))) \delta_i(B(t)) \quad (7.5)$$

where

$$R_j(t, h(t)) = \frac{e^{v_j(t, h(t))}}{\sum_{k \in B(t)} e^{v_k(t, h(t))}}, \quad (7.6)$$

$$\zeta(t-1, t, h(t)) = \frac{\sum_{k \in B(t-1)} e^{v_k(t-1, h(t-1) + \beta(t-1) - \beta(t))}}{\sum_{k \in B(t)} e^{v_k(t, h(t))}}, \quad (7.7)$$

$B(s), B(t) \in \mathfrak{S}$  and  $v_j(t, h(t))$  is defined by (7.3)

A proof of Theorem 9 is given in the appendix.

#### Corollary 6

Under the assumptions of Theorem 9 the c.d.f. of the indirect utility,  $\max_{k \in B} U_k(t), B \in \mathfrak{S}$ , depends on  $\{J(\tau), \tau \leq t\}$  solely through

$$\sum_{k \in B} \exp(v_k(t, h(t))).$$

The result of Corollary 6 follows directly from Corollary 2, (7.1) and (7.2).

Corollary 6 implies that the life cycle consistent property discussed in Section 4 also holds in the case with utilities that are experience-dependent extremal processes, provided the agent does not take into account that current behavior may alter future preferences.

It is important to notice that in contrast to  $P_j(t)$  in (3.9),  $R_j(t, h(t))$  can of course not be interpreted as the marginal choice probability at time  $t$  since it depends on the choice history. It can, however, be interpreted as the marginal choice probability at time  $t$  for an agent equipped with preferences that have been altered by experience.

We may, analogous to Section 5 model state dependence effect in the reparameterized version in which  $\beta(t) = t\theta$ , and  $v_j(t, h(t))$  is substituted by  $w_j(t, h(t))$  defined by

$$e^{w_j(t, h(t))} = e^{v_j(t, h(t))} - e^{v_j(t-1, h(t-1)) - \theta} \quad (7.8)$$

which implies that

$$e^{v_j(t,h(t))} = \sum_{\tau=1}^t e^{w_j(\tau,h(\tau)) + (\tau-t)\theta}. \quad (7.9)$$

From (7.9), (7.4) and (7.6) it follows that  $Q_{ij}(t-1, t, h(t))$  can be expressed as

$$Q_{ij}(t-1, t, h(t)) = \frac{e^{w_j(t,h(t))}}{\sum_{k \in B(t)} \sum_{\tau=1}^t e^{w_k(\tau,h(\tau)) - (t-\tau)\theta}} \quad (7.10)$$

for  $i \neq j$ . The transition probability given a transition has a structure that is completely analogous to (5.5), i.e.,

$$\pi_{ij}(t-1, t, h(t)) = \frac{e^{w_j(t,h(t))}}{\sum_{k \in B(t) \setminus \{i\}} e^{w_k(t,h(t))}}. \quad (7.11)$$

We realize now that both  $\{w_j(t, h(t))\}$  as well as the taste persistent measure  $\theta$  are separately identified. From (7.10) we get that

$$\log \left[ \frac{Q_{ij}(t-1, t, h(t))}{Q_{21}(t-1, t, h(t))} \right] = w_j(t, h(t)) - w_1(t, h(t)). \quad (7.12)$$

Eq. (7.12) means that  $w_j(t, h(t)) - w_1(t, h(t))$  is non-parametrically identified. For example, if  $w_j(t, h(t))$  has the structure

$$w_j(t, h(t)) = \sum_{r=1}^s (\beta_{1r} f_{1r}(Z_j(t)) + \beta_{2r} f_{2r}(Z_j(t), h(t))) \quad (7.13)$$

where  $\{f_{kr}\}$  are known functions and  $\{\beta_{kr}\}$  are unknown parameters,  $k=1,2$ ,  $r=1,2,\dots,s$ , then  $\{\beta_{kr}\}$  are identified under rather general conditions on  $\{f_{kr}(\cdot)\}$ .

Finally, when  $w_j(t, h(t))$  has been determined,  $\theta$  is identified because (7.10) implies that

$$\frac{e^{w_j(t,h(t))}}{Q_{ij}(t-1, t, h(t))} - \sum_{k \in B(t)} e^{w_k(t,h(t))} = e^{-\theta} \cdot \frac{e^{w_j(t-1,h(t-1))}}{Q_{ij}(t-2, t-1, h(t-1))} \quad (7.14)$$

for  $i \neq j$ .

### Example (Heckman, 1981b)

Consider the labor supply example analyzed by Heckman (1981b). Let  $U_2(t)$  be the utility of working and  $U_1(t)$  the utility of not working. If we assume that the transition probabilities given by (7.10) are specified as

$$w_1(t, h(t)) = w_1(t) = Z_1(t) \alpha_1 \quad (7.15)$$

and

$$w_2(t, h(t)) = Z_2(t) \alpha_2 + \delta D(t-1), \quad (7.16)$$

where  $\theta > 0$ ,  $Z_1(t)$  is a vector that may consist of age, length of schooling and number of small children,  $Z_2(t)$  may be some function of the marginal wage rate (or instruments for the marginal wage rate),  $D(\tau)$  is equal to one if the agent has worked in period  $\tau$  and zero otherwise,  $\alpha_1$ ,  $\alpha_2$  and  $\delta$  are parameters to be estimated. In the formulation above  $\delta = 0$  implies PTPP, otherwise there is state dependence in that the agents utility for work is affected by work experience.

Clearly, this model is identified and the specification (7.15) and (7.16) can be exploited to form the likelihood for a sample of individual work histories to estimate  $\alpha_1$ ,  $\alpha_2$ , the taste persistence parameter  $\theta$  and the state dependence parameter  $\delta$ .<sup>9</sup>

## **8. Conclusions**

In this paper we have considered the problem of functional form and stochastic structure in intertemporal discrete choice models.

It is demonstrated that a particular extension of Luce IIA axiom implies a random utility model where the utilities are extremal processes. When the choice set process is non-decreasing this model has the Markov property with a particular structure of the transition probabilities. It is also demonstrated that this model is, under specific assumptions, consistent with optimizing behavior in a life cycle context where the (chosen) expenditure path can be treated as if it consisted of exogenous explanatory variables in the probabilities that correspond to the discrete choices.

Finally, we discuss how the choice model can be extended to allow for time varying choice sets and structural state dependence. In the case with time varying choice sets it turns out that the model structure is not necessarily markovian. Specifically, the transition probabilities in this case are shown to depend on past choices in a particular way. This property is of considerable interest for the ability to distinguish between taste persistence and structural state dependence because it demonstrates that dependence on past choices can arise solely as a result from (exogenous) variations in the choice constraints.

The framework developed in this paper is analytically tractable and it therefore appears convenient for empirical applications.

## Footnotes

- 1 Loosely speaking, a separable function is, by definition, in a certain sense determined by its values in an everywhere-dense, enumerable set of points. A stochastic process is called separable when its sample functions process, with probability one, have this property. However, for a rigorous definition of separability the reader is referred to any book on the theory of probability and stochastic processes. Recall that a process  $\{Y(t), t > 0\}$  being continuous in probability means that  $\exists \delta > 0$  such that for any  $\eta_1, \eta_2 > 0$

$$P(\|Y(t) - Y(s)\| > \eta_1) < \eta_2$$

whenever  $|t - s| < \delta$ , where  $\|\cdot\|$  is the standard Euclidian metric.

- 2 In Dagsvik (1983), p. 11, a particular behavioral assumption (Axiom 1) was proposed. This axiom is weaker than A1. Axiom 1 makes a statement analogous to IIA for choice careers  $J(t_1) = j_1, J(t_2) = j_2, J(t_3) = j_3$  at time epochs  $t_1 < t_2 < t_3$ , where  $j_1, j_2$  and  $j_3$  are all different.
- 3 Heckman (1981a) calls PTPP "habit persistence". We prefer however the notion PTPP since habit persistence may yield association to dependence on past choice experience.
- 4 In general it appears to be rather difficult to extend the static random utility model to the intertemporal setting when we allow for serial correlation in the utilities. Several authors (see Mortensen (1986), Olsen et al. (1986) and Rust (1987)) specify Markovian random utility models that are consistent with choice under uncertainty about the environment. However, in these models the utilities are serially independent. Heckman (1978, 1981a) considers models where the utilities are Gaussian processes. Unfortunately, in the general case with more than two alternatives, serially correlated Gaussian utility processes yield intractable choice probabilities. In particular, a Markovian choice model cannot be obtained as a special case unless the (Gaussian) utilities (for a given alternative) are serially independent (which is only possible in discrete time). McFadden (1984) considers a dynamic model for binary choice based on generalized extreme value distributions.
- 5 Recall that the type III extreme value distribution has the form  $\exp(-ae^{-bx})$  where  $a > 0, b > 0$  are constants (cf. Resnick, 1987). There is a close link between type III, type II and type I extreme value distributions. For example, the type I c.d.f. has the form  $\exp(-ax^{-b})$  with  $a > 0, b > 0$  and it is related to the type III c.d.f. in the following manner: If a random variable  $Y$  has type I c.d.f. then  $\log Y$  has type III c.d.f. Thus, in the context of characterizing the properties of random utility functions type I and type III c.d.f. are equivalent since the mapping  $y \rightarrow \log y$  is increasing. It can easily be demonstrated that if "location shift" (cf. Assumption A4) is substituted by "a multiplicative positive constant", then we must substitute "type III" by "type I" in Theorem 2.
- 6 See Lindberg et al. (1994) for a correction of the proof of the main result in Strauss (1979), and Robertson and Strauss (1981).

- 7 Note that we may interpret  $e^{w_j(t)}$  and  $e^{v_j(t)}$  as mean utilities. Specifically, consider the utility function

$$U_j^* = e^{w_j + \kappa \epsilon_j} / \Gamma(1 - \kappa)$$

where  $P(\epsilon_j \leq x) = \exp(-e^{-x})$  and  $\kappa$  is a positive constant less than one. The utility  $U_j^*$  is of course equivalent to  $w_j + \kappa \epsilon_j$ . The mean of  $U_j^*$  is given by

$$E U_j^* = e^{w_j}.$$

- 8 Similarly to Dagsvik (1983) it can also be proved that  $\zeta(s,t)$  has the interpretation as

$$\zeta(s,t) = \rho \left( \text{corr} \left\{ \max_{k \in B(s)} U_k(s), \max_{k \in B(t)} U_k(t) \right\} \right)$$

where  $\rho : [0,1] \rightarrow [0,1]$  is an increasing function with  $\rho(0) = 0$  and  $\rho(1) = 1$ . The function  $\rho(\cdot)$  has the form

$$\rho(y) = \frac{6}{\pi^2} \int_0^y \frac{\log x \, dx}{1-x}$$

(cf. Tiago de Oliveira, 1973).

- 9 It is understood here that the sample is homogeneous in the sense that  $\alpha_1$ ,  $\alpha_2$ ,  $\theta$  and  $\delta$  do not vary across individuals.

It is interesting that in Heckman's (1981b) analysis, where he postulates a discrete time probit model with general serial dependence in the utilities and with experience dependency, the Markovian structure of the utilities emerges (conditional on experience) as a result from the empirical analysis. However, as mentioned above, continuous time Gaussian utility processes cannot generate Markovian choice models in contrast to the choice models that follow from extremal utility processes. In discrete time the corresponding Probit choice model is Markovian only when the Gaussian utility processes have zero autocorrelation.

## Appendix

### Lemma 1

Let  $F(x,y)$  be a bivariate (type III) extreme value distribution. Then  $-\log F(-x,-y)$  is convex. If  $F(x,y)$  is continuous the left and right derivatives,  $\partial F_{\pm}(x,y)/\partial x$  and  $\partial F_{\pm}(x,y)/\partial y$ , exist and are non-decreasing.

### Proof:

Let  $L(x,y) = -\log F(-x,-y)$ . Since  $F$  is a c.d.f. it follows that  $L$  is non-decreasing. Moreover, since  $F$  is a bivariate extreme value distribution it follows by Proposition 5.11, p. 272 in Resnick (1987) that there exists a finite measure  $\mu$  on

$$\Delta = \{z \in \mathbb{R}_+^2 : z_1^2 + z_2^2 = 1\}$$

such that

$$L(x,y) = \int \max(z_1 e^x, z_2 e^y) \mu(dx, dy).$$

Since  $z_1 e^x$  and  $z_2 e^y$  are convex functions it follows that  $L(x,y)$  is convex. Since  $L(x,y)$  is convex the left and right derivatives of  $F(x,y)$  exist. (See for example Kawata, Theorem 1.11.1 p. 27.)

Q.E.D.

### Proof of Theorem 3:

We assume  $n = 2$  since the general case is completely analogous. As in eq. (4.8) let  $B(s) = B(t-) = \{1,2\}$ ,  $s < t$ , and  $B(t) = \{1,2,3\}$ .

Consider the case with choices at two moments in time,  $s$  and  $t$ . Let  $F_j(s,t;x,y)$  be the c.d.f. of  $(U_j(s), U_j(t))$ . Note that since  $\{U_j(t), t \geq 0\}$  is assumed continuous in probability it follows that  $F_j(s,t; x,y)$  is continuous in  $(x,y)$ . Let

$$G_i(x,y) = -\log F_i(s,t;x,y).$$

We have, since the utility processes are continuous in probability, that

$$\begin{aligned} & P(U_1(s) > U_2(s), U_2(t-) > U_1(t-), U_3(t) > \max(U_1(t), U_2(t))) = P(U_1(s) > U_2(s), U_3(t) > U_2(t) > U_1(t)) \\ & = \int_{\substack{y_1 < y_2 < y_3 \\ x_1 > x_2}} F_1(s,t; dx_1, dy_1) F_2(s,t; dx_2, dy_2) F_3(s,t; \infty, dy_3) = \int (1 - F_3(s,t; \infty, y)) F_2(s,t; x, dy) F_1(s,t; dx, y). \end{aligned} \tag{A.1}$$

Since the marginal distribution of  $U_j(t)$  is type III extreme value (Theorem 2) we can write  $F_j(s,t;\infty,y)$  as

$$F_j(s,t;\infty,y) = \exp(-m_j e^{-y}) \quad (\text{A.2})$$

where  $\log m_j = EU_j(t) - 0.5772$ . By Lemma 1 the first order left and right derivatives of  $G_i(x,y)$  exist. From now on we shall use the notion "derivative of  $G_i(x,y)$ " meaning the respective (first order) right derivative. Since F by A4 is a multivariate extreme value distribution it follows that for  $z \in \mathbb{R}$

$$G_i(x,y) = e^{-z} G_i(x-z, y-z).$$

Hence

$$F_2(s,t;x,dy)F_1(s,t;dx,y) = \left[ \exp(-e^{-y} (G_1(x-y,0) + G_2(x-y,0))) \right] \cdot e^{-2y} \partial_1 G_1(x-y,0) \partial_2 G_2(x-y,0) dx dy \quad (\text{A.3})$$

where  $\partial_j$  denotes the partial derivative with respect to component  $j$ . Let  $h_i(x) = G_i(-x,0)$ . From the relationship

$$G_i(x,y) = e^{-y} G_i(x-y,0) = e^{-y} h_i(y-x)$$

it follows that

$$\partial_2 G_i(-x,0) = -h_i(x) + h_i'(x). \quad (\text{A.4})$$

By Lemma 1,  $h_i(x)$  is convex and therefore has derivatives that are non-decreasing. From (A.1), (A.2), (A.3) and (A.4) it follows after the change of variable  $x = -u + y$ , that

$$\begin{aligned} & P(U_1(s) > U_2(s), U_3(t) > U_2(t) > U_1(t)) \\ &= \int_{\mathbb{R}^2} (1 - \exp(-m_3 e^{-y})) \left[ \exp(-e^{-y} (h_1(u) + h_2(u))) \right] h_1'(u) (h_2(u) - h_2'(u)) e^{-2y} du dy \quad (\text{A.5}) \\ &= \int_{\mathbb{R}} \frac{h_1'(u) (h_2(u) - h_2'(u)) du}{h(u)^2} - \int_{\mathbb{R}} \frac{h_1'(u) (h_2(u) - h_2'(u)) du}{(h(u) + m_3)^2}, \end{aligned}$$

where  $h(u) = h_1(u) + h_2(u)$ . Now Assumption A1 implies that

$$\begin{aligned} & P(U_1(s) > U_2(s), U_3(t) > U_2(t) > U_1(t)) \\ &= P(U_3(t) > \max(U_1(t), U_2(t))) P(U_1(s) > U_2(s), U_2(t) > U_1(t)) \quad (\text{A.6}) \\ &= \frac{m_3}{m_1 + m_2 + m_3} \cdot P(U_1(s) > U_2(s), U_1(t) < U_2(t)). \end{aligned}$$

From (A.5) we obtain, by letting  $m_3 \rightarrow \infty$ , that

$$P(U_1(s) > U_2(s), U_1(t) < U_2(t)) = \int_{\mathbb{R}} \frac{h'_1(u)(h_2(u) - h'_2(u))du}{h(u)^2}. \quad (\text{A.7})$$

Hence, (A.5), (A.6) and (A.7) imply that

$$\frac{m_1 + m_2}{m_1 + m_2 + m_3} \int_{\mathbb{R}} \frac{h'_1(u)(h_2(u) - h'_2(u))du}{h(u)^2} = \int_{\mathbb{R}} \frac{h'_1(u)(h_2(u) - h'_2(u))du}{(h(u) + m_3)^2}. \quad (\text{A.8})$$

Suppose now that  $x = r \geq -\infty$  is the largest point at which  $h'_1(x) + h'_2(x) = 0$ . Then since  $h'_1(x)$  is nondecreasing it must be true that  $h'_1(x) = h'_2(x) = 0$  for  $x \leq r$ . As a consequence the mapping  $\psi: \mathbb{R}_+ \rightarrow [r, \infty)$  defined by

$$z = h(\psi(z)) - h(0) \quad (\text{A.9})$$

exists, is invertible and has (right) derivative everywhere on  $\mathbb{R}_+$ . By change of variable

$$u \rightarrow \psi^{-1}(u) = z$$

(A.8) takes the form

$$\frac{m_1 + m_2}{m_1 + m_2 + m_3} \int_q^{\infty} \frac{f(z) dz}{(h(0) + z)^2} = \int_q^{\infty} \frac{f(z) dz}{(h(0) + z + m_3)^2}, \quad (\text{A.10})$$

where

$$q = h(-\infty) - h(0) = h(r) - h(0),$$

and

$$f(z) = h'_2(\psi(z))\psi'(z)(h_1(\psi(z)) - h'_1(\psi(z))). \quad (\text{A.11})$$

But the right hand side of (A.10) is a generalized Stieltjes transform of  $f(\cdot)$  (see Widder (1941)), evaluated at  $h(0) + m_3$ . Due to the uniqueness property of the generalized Stieltjes transform (A.10) implies that  $f(\cdot)$  must be constant for  $z \geq q$ , since the left hand side of (A.10) is the generalized Stieltjes transform of a constant. As a consequence, we must have that  $m_1 + m_2 = h(0)$ . From the definition of  $\psi(z)$  we get

$$1 = h'(\psi(z))\psi'(z). \quad (\text{A.12})$$

Hence, (A.11) and (A.12) with  $u = \psi(z)$  yield

$$h'_1(u)(h_2(u) - h'_2(u)) = h'(u)C_{12} \quad (\text{A.13})$$

for  $u > r$ , where  $C_{12}$  is a constant. Similarly we get

$$h'_2(u)(h_1(u) - h'_1(u)) = h'(u)C_{21} \quad (\text{A.14})$$

for  $u > r$ , where  $C_{21}$  is another constant. By subtracting (A.14) from (A.13) we get

$$h'(u)h_2(u) - h'_2(u)h(u) = h'(u)(C_{12} - C_{21})$$

which, when dividing by  $h(u)^2$  becomes equal to

$$\frac{h'_2(u)h(u) - h_2(u)h'(u)}{h(u)^2} = \frac{h'(u)(C_{21} - C_{12})}{h(u)^2}. \quad (\text{A.15})$$

Next, integrating both sides of (A.15) yields

$$\frac{h_2(u)}{h(u)} = \frac{C_{12} - C_{21}}{h(u)} + d_1$$

for  $u > r$ , where  $d_1$  is a constant. Hence we obtain

$$h_2(u) = C_{12} - C_{21} + h(u)d_1 \quad (\text{A.16})$$

for  $u > r$ . By inserting (A.16) into (A.14) we get

$$h'(u)(h_1(u) - h'_1(u))d_1 = h'(u)C_{21}$$

for  $u > r$ , which is equivalent to

$$h_1(u) - h'_1(u) = C_{21}/d_1. \quad (\text{A.17a})$$

Similarly, it follows that

$$h_2(u) - h'_2(u) = C_{12}/d_2 \quad (\text{A.17b})$$

for  $u > r$ . Eq. (A.17) is a first order differential equation which has a solution of the form

$$h_j(u) = \alpha_j + \beta_j e^u \quad (\text{A.18a})$$

for  $u > r \geq -\infty$ ,  $j=1,2$ . Since  $h'_j(u) = 0$  for  $u \leq r$  and  $h_j(u)$  is continuous we get from (A.18a) that

$$h_j(u) = \alpha_j + \beta_j e^r \quad (\text{A.18b})$$

for  $u \leq r$ . As a consequence

$$G_j(x, y) = e^{-y} h_j(y - x) = \alpha_j e^{-y} + \beta_j \exp(-\min(x, y - r)). \quad (\text{A.19})$$

From (A.19) we obtain that for  $s < t$

$$P(U_j(t) \leq y | U_j(s) = x) = 0 \quad (\text{A.20})$$

when  $y < x + r$ , and

$$P(U_j(t) \leq y | U_j(s) = x) = P(U_j(t) \leq y) \quad (\text{A.21})$$

when  $y > x + r$ . Eq. (A.20) means that  $\{U_i(t)\}$  is non-decreasing with probability one. Eq. (A.21) means that conditional on  $U_i(t) > U_i(s)$  then  $U_i(t)$  is stochastically independent of  $U_i(s)$ . But then we must have that  $\{U_i(t)\}$  is equivalent to the utility process defined by

$$U_i(t) = \max(U_i(s), W_i(s, t)) + r \quad (\text{A.22})$$

where  $W_i(s, t)$  is extreme value distributed and independent of  $U_i(s)$ . Since  $U_1(t) - U_2(t)$  is independent of  $r$  for any  $t$  we may without loss of generality choose  $r = 0$ . But then (A.22) defines the extremal process which was to be proved.

Q.E.D.

### Lemma 2

Suppose  $\{U_k(t), t \geq 0\}$ ,  $k = 1, 2, \dots, m$ , are independent extremal processes. Let

$V(t) = \max_{k \in B(t)} U_k(t)$ . If  $B(s) \subset B(t)$  for  $s < t$ , then there exists a mapping  $f(\cdot)$  such that

$$J(t, B(t)) = f(J(s, B(s)), V(s), \xi(s, t)) \quad (\text{A.23})$$

where  $\xi(s, t)$  is a random vector which is independent of  $(J(s, B(s)), V(s))$ .

### Proof:

The proof is similar to the proof of Theorem 2.1 in Resnick and Roy (1990). Due to (3.11) we can express  $V(t)$  as

$$V(t) = \max(V(s), W(s, t))$$

where  $W(s, t)$  is a random variable that is independent of  $V(s)$ . Hence

$$J(t, B(t)) = J(s, B(s))1_{[W(s,t) < V(s)]} + \sum_i i 1_{[V(s) < W(s,t) = W_i(s,t)]} \equiv f(J(s, B(s)), V(s), \xi(s, t))$$

where

$$\xi(s, t) = (W_1(s, t), W_2(s, t), \dots, W_m(s, t)),$$

which shows that  $J(t, B(t))$  can be expressed as a function of  $(J(s, B(s)), V(s))$  and something that is independent of  $(J(s, B(s)), V(s))$ .

Q.E.D.

### Lemma 3

Let  $C_k, B \in \mathfrak{S}$ ,  $C_k \subset B, k = 1, 2, \dots, K$ , and assume that the utilities are independent extremal processes (with extreme value marginals). Then for  $s < t$

$$\begin{aligned} & P\left(J(t, B) = j, \max_{k \in B} U_k(t) \leq y \mid J(\tau, C_k), \tau \leq s, k = 1, 2, \dots, K\right) \\ &= P\left(J(t, B) = j \mid J(s, C_1)\right) P\left(\max_{k \in B} U_k(t) \leq y\right). \end{aligned} \tag{A.24}$$

### Proof:

Since  $\{J(s, C_k), s > 0\}$  is a Markov chain for each  $k$  it follows that also

$\{J(s, C_1), J(s, C_2), \dots, J(s, C_K), s > 0\}$  is a Markov chain. Thus

$$\begin{aligned} & P\left(J(t, B) = j, \max_{k \in B} U_k(t) \leq y \mid J(\tau, C_k), \tau \leq s, k = 1, 2, \dots, K\right) \\ &= P\left(J(t, B) = j, \max_{k \in B} U_k(t) \leq y \mid J(s, C_k), k = 1, 2, \dots, K\right). \end{aligned}$$

For expository simplicity we shall go through the rest of the proof for the case with  $B = C_1 = \{1, 2, 3\}$ ,  $C_2 = \{1, 2\}$ ,  $C_3 = \{1, 3\}$ ,  $C_4 = \{2, 3\}$ . Then the information about the choices from  $C_1, C_2, C_3, C_4$ , is equivalent to the information about the ranking of the alternatives at time  $s$ . Under the extremal process assumptions we can write

$$\begin{aligned} F_j(x_j, y_j) &\equiv P(U_j(s) \leq x_j, U_j(t) \leq y_j) \\ &= \exp\left(-a_j e^{-\min(x_j, y_j)} - (b_j - a_j) e^{-y_j}\right) \end{aligned} \tag{A.25}$$

where

$$\log a_j = E U_j(s) - 0.5772, \quad \log b_j = E U_j(t) - 0.5772.$$

From (A.25) it follows that

$$\begin{aligned}
& P(\max(U_1(t), U_2(t)) < U_3(t) \leq y, U_2(s) < U_3(s) < U_1(s)) \\
&= \int_{\substack{x_2 < x_3 < x_1 \\ y_1 < y_2, y_2 < y_3, y_3 < y}} F_1(dx_1, dy_1) F_2(dx_2, dy_2) F_3(dx_3, dy_3) \\
&= \int_{\substack{y_3 < y \\ x_2 < x_3 < x_1}} F_1(dx_1, y_3) F_2(dx_2, y_3) F_3(dx_3, dy_3) \\
&= \int_{y_3 < y} (F_1(\infty, y_3) - F_1(x_3, y_3)) F_2(x_3, y_3) F_3(dx_3, dy_3) \\
&= \int_{x_3 < y_3 \leq y} (\exp(-b_1 e^{-y_3}) - \exp(-a_1 e^{-x_3} - (b_1 - a_1) e^{-y_3})) \exp(-a_2 e^{-x_3} - (b_2 - a_2) e^{-y_3}) \\
&\quad \cdot \exp(-a_3 e^{-x_3} - (b_3 - a_3) e^{-y_3}) (b_3 - a_3) a_2 e^{-x_3 - y_3} dx_3 dy_3 \\
&\quad + \int_{x_3 = y_3 < y} (\exp(-b_1 e^{-y_3}) - \exp(-b_1 e^{-y_3})) \exp(-b_2 e^{-y_3}) \exp(-b_3 e^{-y_3}) (b_3 - a_3) e^{-y_3} dy_3 \\
&= \frac{a_1 a_3 (b_3 - a_3) \exp(-(b_1 + b_2 + b_3) e^{-y})}{(a_1 + a_2 + a_3)(a_2 + a_3)(b_1 + b_2 + b_3)}. \tag{A.26}
\end{aligned}$$

Now since

$$P\left(\max_{k \in B} U_k(t) \leq y\right) = \exp(-(b_1 + b_2 + b_3) e^{-y}) \tag{A.27}$$

(A.26) implies that  $\max_{k \in B} U_k(t)$  is independent of  $(J(s, C_1), J(s, C_2), J(s, C_3), J(s, C_4))$ . Moreover, it

follows from Luce and Suppes (1965), p. 354 that

$$P(U_2(s) < U_3(s) < U_1(s)) = \frac{a_1}{a_1 + a_2 + a_3} \cdot \frac{a_3}{a_2 + a_3} \tag{A.28}$$

and from Theorem 4 that

$$P(J(t, B) = 3 | J(s, B) = 1) = \frac{b_3 - a_3}{b_1 + b_2 + b_3}. \tag{A.29}$$

By inserting (A.27), (A.28) and (A.29) into (A.26) the conclusion of the theorem follows in the particular case considered here. In the general case the argument is completely analogous.

Q.E.D.

Lemma 4

Suppose that the utility processes are independent extremal processes (each extreme value marginals). Let  $t_1 < t_2 < \dots < t_n < t$  be arbitrary points in time. If  $B(t_n) = B(t)$  then

$$P\left(J(t, B(t)) = j \mid J(t_k, B(t_k)), k = 1, 2, \dots, n\right) = P\left(J(t, B(t)) = j \mid J(t_n, B(t))\right). \quad (\text{A.30})$$

Proof:

$$\text{Let } Q = \{r: B(t_r) \subset B(t)\}, \quad V_t(\tau) \equiv \max_{k \in B(t)} U_k(\tau)$$

and

$$H \equiv P\left(J(t, B(t)) = j \mid J(t_k, B(t_k)) = i_k, k = 1, 2, \dots, n\right)$$

with  $B(t) = B(t_n)$ . We have, for  $i_k \in B(t_k)$ ,

$$H = E\left\{P\left(J(t, B(t)) = j \mid J(t_r, B(t_r)) = i_r, r \in Q, J(t_k, B(t_k)), V_t(t_k), U_q(t_k), q \in B(t_k) \setminus B(t), k \notin Q\right) \cdot 1_{[J(t_k, B(t_k)) = i_k, k \notin Q]}\right\}.$$

Since  $U_q(t_k)$  is independent of  $V_t(\tau)$ ,  $J(t_k, B(t_k))$ ,  $J(t_r, B(t_r))$ ,  $J(t, B(t))$  for  $q \notin B(t)$ ,  $r \in Q$ ,  $H$  reduces to

$$H = E\left\{P\left(J(t, B(t)) = j \mid J(t_r, B(t_r)) = i_r, r \in Q, J(t_k, B(t_k)), V_t(t_k), k \notin Q\right) 1_{[J(t_k, B(t_k)) = i_k, k \notin Q]}\right\}$$

which by Lemma 2 reduces further to

$$\begin{aligned} H &= E\left\{P\left(J(t, B(t)) = j \mid J(t_n, B(t_n)) = i_n, V_t(t_n)\right) 1_{[J(t_k, B(t_k)) = i_k, k < n]}\right\} \\ &= E\left\{P\left(J(t, B(t)) = j \mid J(t_n, B(t_n)) = i_n, V_t(t), V_t(t_n)\right) 1_{[J(t_k, B(t_k)) = i_k, k < n]}\right\} \\ &= E\left\{P\left(J(t, B(t)) = j \mid J(t_n, B(t_n)) = i_n, V_t(t)\right) 1_{[J(t_k, B(t_k)) = i_k, k < n]}\right\}. \end{aligned} \quad (\text{A.31})$$

Since  $V_t(t)$  is independent of  $(J(t, B(t)), J(t_n, B(t_n)))$ , (A.31) implies that

$$\begin{aligned} H &= E\left\{P\left(J(t, B(t)) = j \mid J(t_n, B(t_n)) = i_n\right) 1_{[J(t_k, B(t_k)) = i_k, k < n]}\right\} \\ &= P\left(J(t, B(t)) = j \mid J(t_n, B(t_n)) = i_n\right). \end{aligned} \quad (\text{A.32})$$

But (A.32) implies that the transition intensities given the choice history only depends on the last choice before time  $t$ .

Q.E.D.

Proof of Theorem 6:

Let  $B(t) \setminus B(t-) \neq \emptyset$ ,  $j \in B(t) \setminus B(t-)$  and let  $t_1 < t_2 < \dots < t_n < t$  be points in time such that one of the  $t_k$ , say  $t_k^*$ , equals  $r(t)$ . Also let  $B(t_n) = B(t-)$  and  $i_k \in B(t_k)$ . Let

$$N \equiv P\left(J(t, B(t)) = j \mid J(t_k, B(t_k)) = i_k, k \leq n\right).$$

From Lemma 3 it follows that

$$\begin{aligned} N &= E \left\{ P\left(J(t, B(t)) = j \mid J(t_k, B(t_k)), k \leq n\right) 1_{[J(t_k, B(t_k)) = i_k, k \leq n]} \right\} \\ &= E \left\{ P\left(J(t, B(t)) = j \mid J(t_n, B(t_n)) 1_{[J(r(t), B(t)) = i_k^*]} \right) \right\} \\ &= P\left(J(t, B(t)) = j \mid J(r(t), B(t)) = i\right). \end{aligned}$$

Now (6.2) follows from (5.3) by replacing  $B$  by  $B(t)$ . Eq. (6.3) follows from the adding-up conditions.

Q.E.D.

Proof of Theorem 8:

Let  $t_1 < t_2 < \dots < t_n < t$  be arbitrary points in time,  $B(t_n) = B(t-)$  and

$$V_i(\tau) = \max_{k \in B(\tau)} U_k(\tau).$$

Let  $i_k \in B(t_k)$  and choose one of the time epochs, say  $t_k^*$  such that  $s(t) = t_k^*$  with  $i = i_k^*$ . By assumption,  $B(t_n) \setminus B(t) \neq \emptyset$ ,  $i_n \in B(t_n) \setminus B(t)$ . Let  $j \in B(t)$ ,

$$M \equiv P\left(J(t, B(t)) = j \mid J(t_k, B(t_k)) = i_k, k = 1, 2, \dots, n\right)$$

and  $Q = \{r : B(t_k) \in B(t)\}$ . We have that

$$M = E \left\{ P\left(J(t, B(t)) = j \mid J(t_r, B(t_r)), U_q(t_k), V_i(t_r), q \in B(t_r) \setminus B(t), r \notin Q, J(t_k, B(t_k)), k \in Q\right) 1_{[J(t_k, B(t_k)) = i_k, k \leq n]} \right\}.$$

Since  $U_q(t_k)$ , for  $q \in B(t_k) \setminus B(t)$ , is independent of  $J(t, B(t)), J(t_r, B(t)), J(t_k, B(t_k)), k \in Q$ ,  $M$  reduces to

$$\begin{aligned} M &= E \left\{ P \left( J(t, B(t)) = j \mid J(t_r, B(t)), V_t(t_r), r \notin Q, J(t_k, B(t_k)) \right) 1_{[J(t_k, B(t_k)) = i_k, k \leq n]} \right\} \\ &= E \left\{ P \left( J(t, B(t)) = j \mid J(t_n, B(t)), V_t(t_n) \right) 1_{[J(t_k, B(t_k)) = i_k, k \leq n]} \right\}. \end{aligned} \quad (A.33)$$

Since  $\{V_t(\tau), 0 < \tau \leq t\}$  is a Markov process and  $V_t(t)$  is independent of  $J(\tau, B(\tau))$  for  $\tau \leq t, B(\tau) \subset B(t)$  (A.33) reduces further to

$$\begin{aligned} M &= E \left\{ P \left( J(t, B(t)) = j \mid J(t_n, B(t)), V_t(t) \right) 1_{[J(t_k, B(t_k)) = i_k, k \leq n]} \right\} \\ &= E \left\{ P \left( J(t, B(t)) = j \mid J(t_n, B(t)) \right) 1_{[J(t_k, B(t_k)) = i_k, k \leq n]} \right\} \\ &= E \left\{ P \left( J(t, B(t)) = j \mid J(t_n, B(t)) \right) 1_{[J(t_k, B(t_k)) = i_k, k \in Q]} \right\} \\ &= P \left( J(t, B(t)) = j \mid J(t_n, B(t)) = i \right). \end{aligned} \quad (A.34)$$

But then the results of Theorem 8 follow immediately from (5.4) and (5.3) with  $B$  is replaced by  $B(t)$ .  
Q.E.D.

### Proof of Theorem 9

From (7.1) and (7.2) it follow that the conditional distribution of  $U_j(t)$  given  $U_j(t-1)$  and given the choice history can be expressed as

$$P \left( U_j(t) \leq y \mid U_j(t-1) = x, h(t) \right) = \begin{cases} \exp \left( - \exp \left( g_j(t, h(t)) - y \right) \right) & \text{for } y \geq x, \\ 0 & \text{otherwise.} \end{cases} \quad (A.35)$$

For expository simplicity let  $B(t) = S$  for all  $t$ . The proof in the general case is completely analogous.

Eq. (A.35) implies that we may write

$$P \left( U(t) \leq y \mid U(t-1) = x, h(t) \right) = M \left( y - g(t, h(t)) \mid x \right) \quad (A.36)$$

where

$$g(t, h(t)) = (g_1(t, h(t)), g_2(t, h(t)), \dots, g_m(t, h(t)))$$

and

$$M(y|\mathbf{x}) = \begin{cases} \exp\left(-\sum_k e^{-y_k}\right), & y_k \geq x_k, \forall k \leq m, \\ 0 & \text{otherwise.} \end{cases}$$

Let  $j(\tau), \tau = 1, 2, \dots$ , be a sequence of choices and define

$$\Omega(t) = \{(\mathbf{u}(1), \mathbf{u}(2), \dots, \mathbf{u}(t)) : \mathbf{u}_{j(\tau)}(\tau) = \max_k \mathbf{u}_k(\tau), \tau \leq t\}.$$

We have

$$P(J(\tau) = j(\tau), \tau = 1, 2, \dots, t) = \int_{\Omega(t)} \prod_{\tau=1}^t dM(\mathbf{u}(\tau) - \mathbf{g}(\tau, \mathbf{h}^*(\tau)) | \mathbf{u}(\tau - t)) \quad (\text{A.37})$$

where  $\mathbf{u}(0) = -\infty$ , and  $\{\mathbf{h}^*(\tau)\}$  are the choice histories

$$\mathbf{h}^*(\tau) = \{J(s) = j(s), s \leq \tau - 1\}.$$

Let  $g_j(\tau), \tau = 1, 2, \dots$ , denote the parameter of the c.d.f. (7.2) under pure taste persistence, and define

$$L(\mathbf{h}^*(t+1), \mathbf{g}(1), \mathbf{g}(2), \dots, \mathbf{g}(t)) \equiv \int_{\Omega(t)} \prod_{\tau=1}^t dM(\mathbf{u}(\tau) - \mathbf{g}(\tau) | \mathbf{u}(\tau - 1)), \quad (\text{A.38})$$

where  $\mathbf{g}(\tau) = (g_1(\tau), g_2(\tau), \dots, g_n(\tau))$ . Clearly, (A.38) is the likelihood function under pure taste persistence. We know from Theorem 4 that the likelihood function (A.38) has a structure that implies a Markovian choice process under pure taste persistence. But when  $(\mathbf{U}(1), \mathbf{U}(2), \dots, \mathbf{U}(t)) \in \Omega(t)$  the choice history, including the choice at time  $t$ , equals  $\mathbf{h}^*(t+1)$  and consequently the terms  $\mathbf{g}(\tau, \mathbf{h}^*(\tau))$ , for  $\tau = 1, 2, \dots$ , remain constant when the integrand in (A.37) is integrated over  $\Omega(t)$ .

Therefore we must have that

$$P(J(\tau) = j(\tau), \tau = 1, 2, \dots, t) = L(\mathbf{h}^*(t+1), \mathbf{g}(1, \mathbf{h}^*(1)), \mathbf{g}(2, \mathbf{h}^*(2)), \dots, \mathbf{g}(t, \mathbf{h}^*(t))). \quad (\text{A.39})$$

Eq. (A.39) implies that conditional on the parameters  $\mathbf{g}(\tau, \mathbf{h}(\tau)), \tau = 1, 2, \dots$ , the choice process is a Markov chain with transition probabilities that have the same structure as in Theorem 1 with  $v_j(t)$  replaced by  $v_j(t, \mathbf{h}(t))$ , defined in (7.3).

Q.E.D.

## References

- Andersen, P.K., L.S. Hansen and N. Keiding (1991): Non- and Semi-parametric Estimation of Transition Probabilities from Censored Observation of a Non-homogeneous Markov Process. *Scandinavian Journal of Statistics*, **18**, 153-167.
- Block, H.D., and J. Marschak (1960: Random Orderings and Stochastic Theories of Response. In I. Olkin, S. Ghurye, W. Hoeffding, W. Madow and H. Mann (eds.), *Contributions to probability and statistics*. Stanford University Press, Stanford, CA.
- Dagsvik, J.K. (1983): Discrete Dynamic Choice: An Extension of the Choice Models of Luce and Thurstone. *Journal of Mathematical Psychology*, **27**, 1-43.
- Dagsvik, J.K. (1988): Markov Chains Generated by Maximizing Components of Multi-dimensional Extremal Processes. *Stochastic Processes and Their Applications*, **28**, 31-45.
- Dagsvik, J.K. (1992): Intertemporal Discrete Choice, Random Tastes and Functional Form. *Discussion Paper*, no. 77, Statistics Norway.
- Dagsvik, J.K. (1995a): The Structure of Intertemporal Models for Myopic Discrete Choice with Random Preferences. *Memorandum 11/95*, Dept. of Economics, University of Oslo.
- Dagsvik, J.K. (1995b): How Large is the Class of Generalized Extreme Value Random Utility Models? *Journal of Mathematical Psychology*, **39**, 90-98.
- de Haan, L. (1984): A Spectral Representation for Max-stable Processes. *Annals of Probability*, **12**, 1194-1204.
- Dwass, M. (1964): Extremal Processes. *Annals of Mathematical Statistics*, **35**, 1718-1725.
- Dwass, M. (1966): Extremal Processes II. *Illinois Journal of Mathematics*, **10**, 381-391.
- Dwass, M. (1974): Extremal Processes III. *Bull. Inst. Math. Acad. Sin.*, **2**, 255-265.
- Fischhoff, B., P. Slovic, and S. Lichtenstein (1980): Knowing What You Want: Measuring Labile Values. In T.S. Wallsten (ed.), *Cognitive Processes in Choice and Decision Behavior*. Erlbaum, Hillsdale, New Jersey.
- Georgescu-Roegen, N. (1958): Threshold in Choice and the Theory of Demand. *Econometrica*, **26**, 157-168.
- Heckman, J.J. (1978): Simple Statistical Models for Discrete Panel Data Developed and Applied to Test the Hypothesis of True State Dependence against the Hypothesis of Spurious State Dependence. *Annales de l'INSEE*, **30-31**, 227-269.
- Heckman, J.J. (1981a): Statistical Models for the Analysis of Discrete Panel Data. In C.F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data*. MIT Press, Cambridge.
- Heckman, J.J. (1981b): Heterogeneity and State Dependence. In S. Rosen (eds.), *Studies in Labor Markets*. Univ. of Chicago Press, Chicago.
- Heckman, J.J. and B. Singer (1985): Social Science Duration Analysis. In J.J. Heckman and B. Singer (eds.), *Longitudinal Analysis of Labor Market Data*, Cambridge University Press, Cambridge.

- Heckman, J.J. (1991): Identifying the Hand of the Past: Distinguishing State Dependence from Heterogeneity. *American Economic Review*, **81**, 75-79.
- Hogarth, R. (1982): *New Directions for Methodology of Social and Behavioral Science: Question Framing and Response Consistency*. Jossey-Bass, San Francisco.
- Kawata, T. (1972): *Fourier analysis in probability theory*. Academic Press, New York.
- Lindberg, P.O., E.A. Eriksson and L.G. Matsson (1995): Invariance of Achieved Utility in Random Utility Models. *Environment and Planning*, **27**, 121-142.
- Luce, R.D. (1959): *Individual Choice Behavior*. Wiley, New York.
- Luce, R.D. and P. Suppes (1965): Preference, Utility, and Subjective Probability. In R.D. Luce, R.R. Bush, and E. Galanter (eds.), *Handbook of Mathematical Psychology*, III, 249-410, Wiley, New York.
- McFadden, D. (1973): Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (eds.) *Frontiers in Econometrics*. Academic Press, New York.
- McFadden, D. (1984): Econometric Analysis of Qualitative Response Models. In Z. Griliches and M.D. Intriligator (eds.), *Handbook of Econometrics*, Vol. II, 1393-1457, North-Holland, Amsterdam.
- Mortensen, D.T. (1986): Job Search and Labor Market Analysis. In O. Ashenfelter and R. Layard (eds.). *Handbook of Labor Economics*, North Holland, Amsterdam.
- Olsen, R.J., D. Alton Smith, and G. Farkas (1986): Structural and Reduced-Form Models of Choice among Alternatives in Continuous Time: Youth Employment under a Guaranteed Jobs Program. *Econometrica*, **54**, 375-394.
- Resnick, S. (1987): *Extreme Value, Regular Variation and Point Processes*. Springer Verlag, New York.
- Resnick, S. and R. Roy (1990): Multivariate Extremal Processes, Leader Processes and Dynamic Choice Models. *Advances in Applied Probability*, **22**, 309-331.
- Robertson, C.A. and D. Strauss (1981): A Characterization Theorem for Random Utility Variables. *Journal of Mathematical Psychology*, **23**, 184-189.
- Rust, J. (1987): Optimal Replacement of GMC Bus Engines. *Econometrica*, **55**, 999-1033.
- Strauss, D. (1979): Some Results on Random Utility Models. *Journal of Mathematical Psychology*, **17**, 32-35.
- Thurstone, L.L. (1927): A Law of Comparative Judgment. *Psychological Review*, **34**, 272-286.
- Tiago de Oliveira, J. (1968): Extremal Processes: Definitions and Properties. *Publications de l'Institut de Statistique*, Université de Paris, **17**, 25-36.
- Tiago de Oliveira, J. (1973): An Extreme Markovian Stationary Process. *Proceedings of the Fourth Conference in Probability Theory*, Acad. Romania, Brasov, 217-225.
- Tversky, A. (1969): Intransitivity of Preferences. *Psychological Review*, **76**, 31-48.

Tversky, A. and D. Kahneman (1983): Extensional versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment. *Psychological Review*, **90**, 293-315.

Weissman, I. (1975): Extremal Processes Generated by Independent Non-identically Distributed Random Variables. *Annals of Probability*, **3**, 172-177.

Widder, D.V. (1941): *The Laplace Transform*. Princeton University Press, New Jersey.

Yellott, J.I., Jr. (1977): The Relationship between Luce's Choice Axiom, Thurstone's Theory of Comparative Judgment and the Double Exponential Distribution. *Journal of Mathematical Psychology*, **15**, 109-146.

## Issued in the series Discussion Papers

- 42 R. Aaberge, Ø. Kravdal and T. Wennemo (1989): Un-observed Heterogeneity in Models of Marriage Dissolution.
- 43 K.A. Mork, H.T. Mysen and Ø. Olsen (1989): Business Cycles and Oil Price Fluctuations: Some evidence for six OECD countries.
- 44 B. Bye, T. Bye and L. Lorentsen (1989): SIMEN. Studies of Industry, Environment and Energy towards 2000.
- 45 O. Bjerkholt, E. Gjelsvik and Ø. Olsen (1989): Gas Trade and Demand in Northwest Europe: Regulation, Bargaining and Competition.
- 46 L.S. Stambøl and K.Ø. Sørensen (1989): Migration Analysis and Regional Population Projections.
- 47 V. Christiansen (1990): A Note on the Short Run Versus Long Run Welfare Gain from a Tax Reform.
- 48 S. Glomsrød, H. Vennemo and T. Johnsen (1990): Stabilization of Emissions of CO<sub>2</sub>: A Computable General Equilibrium Assessment.
- 49 J. Aasness (1990): Properties of Demand Functions for Linear Consumption Aggregates.
- 50 J.G. de Leon (1990): Empirical EDA Models to Fit and Project Time Series of Age-Specific Mortality Rates.
- 51 J.G. de Leon (1990): Recent Developments in Parity Progression Intensities in Norway. An Analysis Based on Population Register Data
- 52 R. Aaberge and T. Wennemo (1990): Non-Stationary Inflow and Duration of Unemployment
- 53 R. Aaberge, J.K. Dagsvik and S. Strøm (1990): Labor Supply, Income Distribution and Excess Burden of Personal Income Taxation in Sweden
- 54 R. Aaberge, J.K. Dagsvik and S. Strøm (1990): Labor Supply, Income Distribution and Excess Burden of Personal Income Taxation in Norway
- 55 H. Vennemo (1990): Optimal Taxation in Applied General Equilibrium Models Adopting the Armington Assumption
- 56 N.M. Stølen (1990): Is there a NAIRU in Norway?
- 57 Å. Cappelen (1991): Macroeconomic Modelling: The Norwegian Experience
- 58 J.K. Dagsvik and R. Aaberge (1991): Household Production, Consumption and Time Allocation in Peru
- 59 R. Aaberge and J.K. Dagsvik (1991): Inequality in Distribution of Hours of Work and Consumption in Peru
- 60 T.J. Klette (1991): On the Importance of R&D and Ownership for Productivity Growth. Evidence from Norwegian Micro-Data 1976-85
- 61 K.H. Alfsen (1991): Use of Macroeconomic Models in Analysis of Environmental Problems in Norway and Consequences for Environmental Statistics
- 62 H. Vennemo (1991): An Applied General Equilibrium Assessment of the Marginal Cost of Public Funds in Norway
- 63 H. Vennemo (1991): The Marginal Cost of Public Funds: A Comment on the Literature
- 64 A. Brendemoen and H. Vennemo (1991): A climate convention and the Norwegian economy: A CGE assessment
- 65 K.A. Brekke (1991): Net National Product as a Welfare Indicator
- 66 E. Bowitz and E. Storm (1991): Will Restrictive Demand Policy Improve Public Sector Balance?
- 67 Å. Cappelen (1991): MODAG. A Medium Term Macroeconomic Model of the Norwegian Economy
- 68 B. Bye (1992): Modelling Consumers' Energy Demand
- 69 K.H. Alfsen, A. Brendemoen and S. Glomsrød (1992): Benefits of Climate Policies: Some Tentative Calculations
- 70 R. Aaberge, Xiaojie Chen, Jing Li and Xuezheng Li (1992): The Structure of Economic Inequality among Households Living in Urban Sichuan and Liaoning, 1990
- 71 K.H. Alfsen, K.A. Brekke, F. Brunvoll, H. Lurås, K. Nyborg and H.W. Sæbø (1992): Environmental Indicators
- 72 B. Bye and E. Holmøy (1992): Dynamic Equilibrium Adjustments to a Terms of Trade Disturbance
- 73 O. Aukrust (1992): The Scandinavian Contribution to National Accounting
- 74 J. Aasness, E. Eide and T. Skjerpen (1992): A Criminometric Study Using Panel Data and Latent Variables
- 75 R. Aaberge and Xuezheng Li (1992): The Trend in Income Inequality in Urban Sichuan and Liaoning, 1986-1990
- 76 J.K. Dagsvik and S. Strøm (1992): Labor Supply with Non-convex Budget Sets, Hours Restriction and Non-pecuniary Job-attributes
- 77 J.K. Dagsvik (1992): Intertemporal Discrete Choice, Random Tastes and Functional Form
- 78 H. Vennemo (1993): Tax Reforms when Utility is Composed of Additive Functions
- 79 J.K. Dagsvik (1993): Discrete and Continuous Choice, Max-stable Processes and Independence from Irrelevant Attributes
- 80 J.K. Dagsvik (1993): How Large is the Class of Generalized Extreme Value Random Utility Models?
- 81 H. Birkelund, E. Gjelsvik, M. Aaserud (1993): Carbon/energy Taxes and the Energy Market in Western Europe
- 82 E. Bowitz (1993): Unemployment and the Growth in the Number of Recipients of Disability Benefits in Norway
- 83 L. Andreassen (1993): Theoretical and Econometric Modeling of Disequilibrium
- 84 K.A. Brekke (1993): Do Cost-Benefit Analyses favour Environmentalists?
- 85 L. Andreassen (1993): Demographic Forecasting with a Dynamic Stochastic Microsimulation Model
- 86 G.B. Asheim and K.A. Brekke (1993): Sustainability when Resource Management has Stochastic Consequences
- 87 O. Bjerkholt and Yu Zhu (1993): Living Conditions of Urban Chinese Households around 1990
- 88 R. Aaberge (1993): Theoretical Foundations of Lorenz Curve Orderings
- 89 J. Aasness, E. Bjørn and T. Skjerpen (1993): Engel Functions, Panel Data, and Latent Variables - with Detailed Results

- 90 I. Svendsen (1993): Testing the Rational Expectations Hypothesis Using Norwegian Microeconomic Data Testing the REH. Using Norwegian Microeconomic Data
- 91 E. Bowitz, A. Rødseth and E. Storm (1993): Fiscal Expansion, the Budget Deficit and the Economy: Norway 1988-91
- 92 R. Aaberge, U. Colombino and S. Strøm (1993): Labor Supply in Italy
- 93 T.J. Klette (1993): Is Price Equal to Marginal Costs? An Integrated Study of Price-Cost Margins and Scale Economies among Norwegian Manufacturing Establishments 1975-90
- 94 J.K. Dagsvik (1993): Choice Probabilities and Equilibrium Conditions in a Matching Market with Flexible Contracts
- 95 T. Kornstad (1993): Empirical Approaches for Analysing Consumption and Labour Supply in a Life Cycle Perspective
- 96 T. Kornstad (1993): An Empirical Life Cycle Model of Savings, Labour Supply and Consumption without Intertemporal Separability
- 97 S. Kverndokk (1993): Coalitions and Side Payments in International CO<sub>2</sub> Treaties
- 98 T. Eika (1993): Wage Equations in Macro Models. Phillips Curve versus Error Correction Model Determination of Wages in Large-Scale UK Macro Models
- 99 A. Brendemoen and H. Vennemo (1993): The Marginal Cost of Funds in the Presence of External Effects
- 100 K.-G. Lindquist (1993): Empirical Modelling of Norwegian Exports: A Disaggregated Approach
- 101 A.S. Jore, T. Skjerpen and A. Rygh Swensen (1993): Testing for Purchasing Power Parity and Interest Rate Parities on Norwegian Data
- 102 R. Nesbakken and S. Strøm (1993): The Choice of Space Heating System and Energy Consumption in Norwegian Households (Will be issued later)
- 103 A. Aaheim and K. Nyborg (1993): "Green National Product": Good Intentions, Poor Device?
- 104 K.H. Alfsen, H. Birkelund and M. Aaserud (1993): Secondary benefits of the EC Carbon/ Energy Tax
- 105 J. Aasness and B. Holtsmark (1993): Consumer Demand in a General Equilibrium Model for Environmental Analysis
- 106 K.-G. Lindquist (1993): The Existence of Factor Substitution in the Primary Aluminium Industry: A Multivariate Error Correction Approach on Norwegian Panel Data
- 107 S. Kverndokk (1994): Depletion of Fossil Fuels and the Impacts of Global Warming
- 108 K.A. Magnussen (1994): Precautionary Saving and Old-Age Pensions
- 109 F. Johansen (1994): Investment and Financial Constraints: An Empirical Analysis of Norwegian Firms
- 110 K.A. Brekke and P. Børing (1994): The Volatility of Oil Wealth under Uncertainty about Parameter Values
- 111 M.J. Simpson (1994): Foreign Control and Norwegian Manufacturing Performance
- 112 Y. Willassen and T.J. Klette (1994): Correlated Measurement Errors, Bound on Parameters, and a Model of Producer Behavior
- 113 D. Wetterwald (1994): Car ownership and private car use. A microeconomic analysis based on Norwegian data
- 114 K.E. Rosendahl (1994): Does Improved Environmental Policy Enhance Economic Growth? Endogenous Growth Theory Applied to Developing Countries
- 115 L. Andreassen, D. Fredriksen and O. Ljones (1994): The Future Burden of Public Pension Benefits. A Microsimulation Study
- 116 A. Brendemoen (1994): Car Ownership Decisions in Norwegian Households.
- 117 A. Langørgen (1994): A Macromodel of Local Government Spending Behaviour in Norway
- 118 K.A. Brekke (1994): Utilitarianism, Equivalence Scales and Logarithmic Utility
- 119 K.A. Brekke, H. Lurås and K. Nyborg (1994): Sufficient Welfare Indicators: Allowing Disagreement in Evaluations of Social Welfare
- 120 T.J. Klette (1994): R&D, Scope Economies and Company Structure: A "Not-so-Fixed Effect" Model of Plant Performance
- 121 Y. Willassen (1994): A Generalization of Hall's Specification of the Consumption function
- 122 E. Holmøy, T. Hægeland and Ø. Olsen (1994): Effective Rates of Assistance for Norwegian Industries
- 123 K. Mohn (1994): On Equity and Public Pricing in Developing Countries
- 124 J. Aasness, E. Eide and T. Skjerpen (1994): Criminometrics, Latent Variables, Panel Data, and Different Types of Crime
- 125 E. Bjørn and T.J. Klette (1994): Errors in Variables and Panel Data: The Labour Demand Response to Permanent Changes in Output
- 126 I. Svendsen (1994): Do Norwegian Firms Form Extrapolative Expectations?
- 127 T.J. Klette and Z. Griliches (1994): The Inconsistency of Common Scale Estimators when Output Prices are Unobserved and Endogenous
- 128 K.E. Rosendahl (1994): Carbon Taxes and the Petroleum Wealth
- 129 S. Johansen and A. Rygh Swensen (1994): Testing Rational Expectations in Vector Autoregressive Models
- 130 T.J. Klette (1994): Estimating Price-Cost Margins and Scale Economies from a Panel of Microdata
- 131 L. A. Grünfeld (1994): Monetary Aspects of Business Cycles in Norway: An Exploratory Study Based on Historical Data
- 132 K.-G. Lindquist (1994): Testing for Market Power in the Norwegian Primary Aluminium Industry
- 133 T. J. Klette (1994): R&D, Spillovers and Performance among Heterogenous Firms. An Empirical Study Using Microdata
- 134 K.A. Brekke and H.A. Gravningsmyhr (1994): Adjusting NNP for instrumental or defensive expenditures. An analytical approach
- 135 T.O. Thoresen (1995): Distributional and Behavioural Effects of Child Care Subsidies
- 136 T. J. Klette and A. Mathiassen (1995): Job Creation, Job Destruction and Plant Turnover in Norwegian Manufacturing
- 137 K. Nyborg (1995): Project Evaluations and Decision Processes
- 138 L. Andreassen (1995): A Framework for Estimating Disequilibrium Models with Many Markets

- 139 L. Andreassen (1995): Aggregation when Markets do not Clear
- 140 T. Skjerpen (1995): Is there a Business Cycle Component in Norwegian Macroeconomic Quarterly Time Series?
- 141 J.K. Dagsvik (1995): Probabilistic Choice Models for Uncertain Outcomes
- 142 M. Rønsen (1995): Maternal employment in Norway, A parity-specific analysis of the return to full-time and part-time work after birth
- 143 A. Bruvoll, S. Glomsrød and H. Vennemo (1995): The Environmental Drag on Long-term Economic Performance: Evidence from Norway
- 144 T. Bye and T. A. Johnsen (1995): Prospects for a Common, Deregulated Nordic Electricity Market
- 145 B. Bye (1995): A Dynamic Equilibrium Analysis of a Carbon Tax
- 146 T. O. Thoresen (1995): The Distributional Impact of the Norwegian Tax Reform Measured by Disproportionality
- 147 E. Holmøy and T. Hægeland (1995): Effective Rates of Assistance for Norwegian Industries
- 148 J. Aasness, T. Bye and H.T. Mysen (1995): Welfare Effects of Emission Taxes in Norway
- 149 J. Aasness, E. Biørn and Terje Skjerpen (1995): Distribution of Preferences and Measurement Errors in a Disaggregated Expenditure System
- 150 E. Bowitz, T. Fæhn, L. A. Grünfeld and K. Moun (1995): Transitory Adjustment Costs and Long Term Welfare Effects of an EU-membership – The Norwegian Case
- 151 I. Svendsen (1995): Dynamic Modelling of Domestic Prices with Time-varying Elasticities and Rational Expectations
- 152 I. Svendsen (1995): Forward- and Backward Looking Models for Norwegian Export Prices
- 153 A. Langørgen (1995): On the Simultaneous Determination of Current Expenditure, Real Capital, Fee Income, and Public Debt in Norwegian Local Government
- 154 A. Katz and T. Bye(1995): Returns to Publicly Owned Transport Infrastructure Investment. A Cost Function/ Cost Share Approach for Norway, 1971-1991
- 155 K. O. Aarbu (1995): Some Issues About the Norwegian Capital Income Imputation Model
- 156 P. Boug, K. A. Mork and T. Tjemsland (1995): Financial Deregulation and Consumer Behavior: the Norwegian Experience
- 157 B. E. Naug and R. Nymoen (1995): Import Price Formation and Pricing to Market: A Test on Norwegian Data
- 158 R. Aaberge (1995): Choosing Measures of Inequality for Empirical Applications.
- 159 T. J. Klette and S. E. Førre (1995): Innovation and Job Creation in a Small Open Economy: Evidence from Norwegian Manufacturing Plants 1982-92
- 160 S. Holden, D. Kolsrud and B. Vikøren (1995): Noisy Signals in Target Zone Regimes: Theory and Monte Carlo Experiments
- 161 T. Hægeland (1996): Monopolistic Competition, Resource Allocation and the Effects of Industrial Policy
- 162 S. Grepperud (1996): Poverty, Land Degradation and Climatic Uncertainty
- 163 S. Grepperud (1996): Soil Conservation as an Investment in Land
- 164 K. A. Brekke, V. Iversen and J. Aune (1996): Soil Wealth in Tanzania
- 165 J. K. Dagsvik, D.G. Wetterwald and R. Aaberge (1996): Potential Demand for Alternative Fuel Vehicles
- 166 J.K. Dagsvik (1996): Consumer Demand with Unobservable Product Attributes. Part I: Theory
- 167 J.K. Dagsvik (1996): Consumer Demand with Unobservable Product Attributes. Part II: Inference
- 168 R. Aaberge, A. Björklund, M. Jäntti, M. Palme, P. J. Pedersen, N. Smith and T. Wennemo (1996): Income Inequality and Income Mobility in the Scandinavian Countries Compared to the United States
- 169 K. Nyborg (1996): Some Norwegian Politicians' Use of Cost-Benefit Analysis
- 170 E. Berg, S. Kverndokk and K. E. Rosendahl (1996): Market Power, International CO<sub>2</sub> Taxation and Petroleum Wealth
- 171 Rolf Aaberge, Ugo Colombino and Steinar Strøm (1996): Welfare Effects of Proportional Taxation: Empirical Evidence from Italy, Norway and Sweden
- 172 J.K. Dagsvik (1996): Dynamic Choice, Multistate Duration Models and Stochastic Structure

Statistics Norway  
Research Department  
P.O.B. 8131 Dep.  
N-0033 Oslo

Tel.: + 47 - 22 86 45 00  
Fax: + 47 - 22 11 12 38

ISSN 0803-074X



**Statistics Norway**  
Research Department