MEMORANDUM

No 06/2006

Justifying Functional Forms in Models for Transitions between Discrete States, with Particular Reference to Employment-Unemployment Dynamics

John K. Dagsvik

ISSN: 0809-8786

Department of Economics
University of Oslo

This series is published by the
University of Oslo
Department of Economics

P. O.Box 1095 Blindern
N-0317 OSLO Norway
Telephone: + 47 22855127
Fax: + 47 22855035
Internet: http://www.oekonomi.uio.no/
e-mail: econdep@econ.uio.no

In co-operation with
The Frisch Centre for Economic Research

Gaustadalleén 21
N-0371 OSLO Norway
Telephone: +47 22 95 88 20
Fax: +47 22 95 88 25
Internet: http://www.frisch.uio.no/
e-mail: frisch@frisch.uio.no

List of the last 10 Memoranda:

No 05	Hild-Marte Bjørnsen and Erik Biørn The Joint Labour Decisions of Farm Couples:A Censored Response Analysis of On-farm and Off-farm Work. 50 pp.
No 04	Morten Henningsen Moving between Welfare Payments. The Case of Sickness Insurance for the Unemployed. 35 pp.
No 03	Rolf Golombek and Michael Hoel Endogenous technology and tradable emission quotas. 16 pp.
No 02	Jo Thori Lind Do the rich vote Conservative because they are rich?. 46 pp.
No 01	Erik Hernæs, John Piggott, Tao Zhang and Steinar Strøm The Determinants of Occupational Pensions. 32 pp.
No 35	Tor Eriksson, Bernt Bratsberg and Oddbjørn Raaum Earnings persistence across generations:Transmission through health? 26 pp.
No 34	Markus Jäntti, Bernt Bratsberg, Knut Røed, Oddbjørn Raaum, Robin Naylor, Eva Österbacka, Anders Björklund and Tor Eriksson American exceptionalism in a new light: a comparison of intergenerational earnings mobility in the Nordic countries, the United Kingdom and the United States. 44 pp.
No 33	Dag Morten Dalen, Tonje Haabeth and Steinar Strøm Price regulation and generic competition in the pharmaceutical market. 44 pp.
No 32	Hilde C. Bjørnland and Håvard Hungnes The commodity currency puzzle. 28 pp.
No 31	Hilde C. Bjørnland Monetary policy and exchange rate interactions in a small open economy. 24 pp.

A complete list of this memo-series is available in a PDF® format at:
http://www.oekonomi.uio.no/memo/

(x:\502\div-gaf\asd\dax\char-func-form-jan06.doc) 25 April 2006

Justifying Functional Forms in Models for Transitions between Discrete States, with Particular Reference to Employment-Unemployment Dynamics

by

John K. Dagsvik

Research Department, Statistics Norway and the Ragnar Frisch Centre for Economic Research

Abstract

This paper proposes a particular axiomatic approach to motivate the choice of functional forms and distribution of unobservables in continuous time models for discrete panel data analysis. We discuss in particular applications with data on transitions between employment and unemployment. This framework yields a characterization of transition probabilities and duration distributions in terms of structural parameters of the utility function and choice constraints. Moreover, it is discussed how the modeling framework can be extended to allow for involuntary transitions, structural state dependence and random effects.

Acknowledgement: This research is part of the project “Mobilizing labour force participation”, conducted at the Frisch Centre for Economic Research and financed by the Norwegian Research Council. I am grateful for useful comments by Knut Røed and Steinar Strøm.

1. Introduction

A difficult challenge in the modeling of behavioral relations is how to justify the choice of functional form and distributional properties of unobservables in these relations. The traditional “solution” to this challenge is to select a flexible class of parametric or semi-parametric specifications purely on grounds of convenience, and then proceed by using statistical inference methods to select the suitable specification within the a priori selected class of specifications. From a scientific perspective the problem with this approach is that one simply does not know how the ad hoc restrictions on the class of specifications affect the analysis. This is due to the fact that it is insufficient to rely solely on statistical inference theory as a strategy to determine functional form and distributional properties of the error terms of structural models. The reason for this is that the class of possible model specifications is very large. Without theoretically justified principles for restricting the empirical specifications it is hard to pin down the correct specification due to measurement error, limited amount of data and non-systematic variation steming from unobserved heterogeneity and bounded rationality. A full nonparametric approach is not possible in practice, because it requires that the analyst has access to an unlimited set of data covering behavioral responses that correspond to every possible and relevant policy regime.

A particularly important version of the functional form challenge arises in the context of dynamic modeling. In the analysis of dynamic behavioral processes it is crucial to identify individual behavioral relations (including the corresponding random error terms), versus properties of the distribution of preferences and choice constraints across the population. This identification is of course fundamental for establishing a modeling framework that can be applied for analyzing structural state dependence effects. More precisely, in some analyses it may be of interest to separate the genuine effect on preferences (or choice constraints) from past choice experience on one hand, from the effect of serial correlation in the unobservables on the other. It is known that without a priori theoretical restrictions on the model this identification problem cannot be settled by statistical methods alone (Heckman, 1991).

From a theoretical point of view it is necessary to clarify what it is meant by "randomness" at the individual level, as perceived by the individual agent. Otherwise one will not be able to distinguish conceptually unobserved heterogeneity from "genuine" randomness. For example, in models for unemployment duration based on search theory the randomness is associated with individual uncertainty with respect to arriving job offers. This includes actual arrivals of job offers as well as the updating of information regarding future uncertain events. However, preferences are usually assumed to be deterministic. In contrast, psychologists have typically found that preferences may be random to the agent himself, cf. Thurstone (1927) and Tversky (1969). The rationale for this is that the agent is viewed as boundedly rational in the sense that he/she may assess different values to the same, or seemingly equivalent alternative, at different points in time.

This paper discusses a particular approach to modeling transitions between labor market states with particular reference to transitions between the two states; “employment” and “unemployment”. Our point of departure is a random utility framework in which the agent, at any point in time goes to the state that yields the highest utility[1]. We view randomness at the individual level as steming from both uncertainty with respect to choice opportunities as well as bounded rationality in the sense described above. Our purpose is to obtain a characterization of the functional form of the model in this setting by means of an axiomatic approach with particular reference to the model structure in the special case with no state dependence. The approach taken in this paper rests critically on a particular definition of absence of state dependence (Axiom 2), and it can be summarized as follows: In the “reference” case of absence of state dependence, the indirect utility (conditional on individual characteristics) is postulated to be independent of current and previously chosen alternatives (states). The intuition of this assumption is that since the utility function represents the value of the respective states, once maximum utility has been achieved, it should be irrelevant for the level of the highest utility which of the states that yields maximum utility. If this were not so, it would mean that not all relevant aspects of the states were captured by the utility function since knowledge about the actual choice and the choice history would represent relevant information about the utility of the current choice. But this contradicts the notion of utility as an “ideal” index which is supposed to capture the value of all relevant aspects of the states.

Under suitable regularity assumptions, it follows from this particular formalization of the notion of absence of state dependence (Axiom 2) that, in this reference case, the utility of being in a given state, viewed as a stochastic process in time, becomes a so called *Extremal process* with deterministic drift. If the utility $U_j(t)$ is an extremal process (with drift) in discrete time it has the property that $U_j(t) = \max(U_j(t-1) - \theta, W_j(t))$, where θ is a non-negative constant and $W_j(t)$, t =1, 2,…, are i.i.d. random variables. The representation of the continuous time version is similar. Apart from a deterministic trend, the extremal process in continuous time changes only through positive jumps of stochastic size that occur at stochastic moments in time. When the utilities of the alternatives are

[1] It is implicit in this formulation that the agent is viewed as always facing job opportunities. This implies no essential restriction because the corresponding utilities of the available jobs may be so low that the jobs are of no interest whatsoever. For example, a job that yields very little income, or is located in another country, may be regarded as uninteresting.

In this paper we do not consider the case in which institutional regulations determined by the employment offices severely restrict the worker’s control of his reservation wage. In such cases, unless the worker (or worker’s household) fulfills specific criteria, he is obliged to accept a job offered by the employment office. This means that the respective managers at the employment offices control to a large extent the arrival process of the individual job offers, since they decide on the basis of the workers formal qualifications which jobs are “suitable”. If the worker declines a job offer that is viewed as suitable by the manager in charge he will normally lose his unemployment benefits, unless he fulfills specific criteria.

extremal processes Dagsvik (1988, 2002) has proved that the choice of state becomes a Markov chain in continuous time. Under the absence of state dependence the Markov property seems intuitive because under reasonable regularity conditions, the weakest possible dependence structure of the choice process in continuous time is represented by the Markov property, with the Bernoulli process as a special limiting case[2]. Second, this assumption implies a particular characterization of the transition intensities (hazard functions) of this Markov chain, in terms of parameters of the underlying utility function. Subsequently, Dagsvik (1998) has demonstrated how one can extend this basic framework to accommodate state dependence and random effects.

The implication that the utilities are extremal processes may at first glance seem counterintuitive. However, the extremal process utility representation is in fact consistent, to some extent, with an old concept in psychology and psychophysics known as *sensory threshold*, introduced by the philosopher Herbart (1824), see Gescheider (1997). By sensory threshold it is meant that mental events have to be stronger than some critical amount in order to be consciously experienced. In our context the random variable $W_j(t)$ may be interpreted as the value of the incoming stimulus that is relevant for the value of alternative j. The special case with $\theta = 0$, corresponds to a particular situation with "perfect" memory in the sense that previous utility evaluations are retained fixed in the agent's consciousness. According to the updating relation above the agent will in this case change his utility of alternative j if the value of the incoming stimulus is greater than the value of the utility in the previous period, whereas otherwise he will not change his assessment of alternative j. In other words, an agent with this type of perfect memory with keep his utility unchanged until sufficiently large stimuli arrive. Thus, new information about negative aspects of the respective alternatives is not taken into account in this case. The case with positive θ corresponds to the case with "imperfect" memory in the sense that previous utility evaluations are depreciated and, as a result, the current arriving stimulus will be taken into account provided $W_j(t) > U_j(t-1) - \theta$. Although the interpretation outlined above is interesting, it is by no means crucial for the theoretical justification of our approach. The rationale for our approach rests fundamentally on the intuition associated with the Axiom 2 described above.

Another feature of the approach taken in this paper regards the characterization of the utility of being unemployed. Many researchers use search theory à la Lippman and McCall (1981) to derive structural restrictions in this context. However, typical formulations of search theoretic models are controversial because there is no generally accepted way of representing this theory explicitly in quantitative empirical relations. In other words, this is yet another example of the insufficiency of economic theory in providing guidance for functional form specifications in empirical relations. In this paper we propose an alternative approach. In this approach the agent is viewed as boundedly rational

[2] In a Bernoulli process the realizations at different points in time are independent.

in the sense that he does not react to a stimulus unless the intensity of the stimulus is higher than some threshold, which is consistent with the idea of sensory threshold mentioned above. A key assumption in this alternative approach is a particular invariance assumption. This assumption is summarized as follows: Suppose first that the arrival rate of job offers change, but that this change is so small that some (boundedly rational) workers do not change their utility of unemployment. Recall that this is possible within our setting with utilities that are extremal processes. Suppose next that the fraction of workers that do not change their utility of unemployment when the arrival rate of job offers changes from λ_1 to λ_2 (say) is less than the fraction of workers that do not change their utility of unemployment when the arrival rate of job offers changes from λ_3 to λ_4 (say). Then the assumption states that this inequality still holds if all arrival rates are multiplied by a common arbitrary positive constant. In other words, this assumption states that as long as the relative change in the respective arrival rates remain unchanged, then the respective fractions of workers that do not change their utility of unemployment may alter, but not to the extent that the inequality is reversed. Analogous invariance postulates have become increasingly common in psychophysics and measurement theories, and in fact also in physics, see for example Falmagne (1985), Luce (1996) and Dagsvik et al. (2006). The rationale for some of these postulates within measurement theory is that psychologists in many studies have found that people seem not to be so responsive to changes in levels as long as relative changes remain constant.

The paper is organized as follows: in sections 2 and 3 we propose a behavioral characterization and interpretation of the utility function as a random function of time. In section 4 we derive implications for the corresponding probabilistic choice model. In section 5 we apply a particular invariance assumption to characterize the utility of unemployment as a function of the arrival rate of job offers. In section 6 we extend the framework to allow for involuntary transitions into unemployment and we derive the corresponding choice model for this case. In section 7 we discuss identification. In section 8 we discuss how one can allow for structural state dependence and in section 9 we discuss briefly examples of representing unobserved heterogeneity in parameters by random effects. In section 10 we consider briefly the example of empirical specification of the hazard function and likelihood functions.

2. Characterization of preferences

To fix ideas we consider a labor market with two states, “employed” (1) and “unemployed” (0). However, the framework we discuss will apply to other applications as well. Examples of leisure activities are different hobbies, sports activities, reading and studies. The individual’s preferences are represented by a utility function that has a particular structure. Let $V_j(t)$ denote the utility to the agent of being in state *j* at age *t* (discrete). This utility function can be thought of as a value function representing, current and future uncertain prospects. Let $J(t) \in \{0,1\}$ denote the state the agent is in at time *t*, i.e., $J(t) = 1$ if the worker is employed time *t*, and $J(t) = 0$ if the agent is unemployed. Thus, formally, $J(t) = 1$ if $V_1(t) > V_0(t)$, and vice versa. Let a denote a "reference age", meaning the age where the agent starts to be exposed to labor market opportunities. The function $\{J(t), t > a\}$ will be a random process since we shall assume that the utility processes, $\{V_j(t), j = 0,1\}$, are random. By this we mean random in the sense of Thurstone; that is, random to the agent himself (Thurstone, 1927). This is motivated by the results from numerous laboratory-type of experiments in which individuals have been found to make different choices in identical experimental settings. This is explained as an effect of bounded rationality in that the decision-makers have difficulties with assessing a precise and definitive value of the respective alternatives. This is so because; (i) the agent may have taste for variation, (ii) he may find it hard to assess the value of the alternatives because he is unsure about his tastes and his perceptions may be influenced by fluctuating moods and whims, (iii) he may have limited information about the alternatives and he may receive unanticipated information over time. In this setting the randomness of tastes cannot be reduced to one of which the decision-maker is capable of representing uncertainty as if it were a lottery with known lottery outcome probabilities, simply because he is unable to predict the distribution of changes in his own psychological states.

Also additional randomness may be present due to variables that are known to the individual agents but are unobservable to the researcher and thus perceived as random by him.

We shall now make assumptions with the purpose of obtaining a behavioral justification of the quantitative structure of the model.

Axiom 1

In the absence of structural state dependence the utility processes, $\{V_j(t), t > a, j = 0,1\}$ *are independent max-stable processes that are separable and continuous in probability, and with type III extreme value marginal c.d.f.*

Recall that the class of max-stable utility processes has the property that it allows one to apply the maximum operation without "leaving the class". That is, the maximum of independent max-stable processes is also a max-stable process. This is analogous to the class of Gaussian processes, which is closed under aggregation. A max-stable process $\{X(t), t > 0\}$, (say), has the property that for any set of points in time, $t_1 < t_2 < ...t_n$, the vector $\left(X\left(t_1\right), X\left(t_2\right),..., X\left(t_n\right)\right)$ is distributed according to the (type III) multivariate extreme value c.d.f., see Resnick (1987). The type III distributions have (standardized) marginal c.d.f. equal to $\exp\left(-\exp(-x)\right)$ [3]. The max-stable property is of course very convenient when dealing with maximization of random utility functions. It also seems like a rather natural invariance property random utility functions should possess, namely that the structure of the c.d.f. of the utility processes should not depend critically on the level of aggregation of alternatives. Note that aggregation of alternatives corresponds to utility maximization because the utility of a subset of alternatives equals the maximum of the utilities of the alternatives of the subset. As regards the significance of the max-stable property, Dagsvik (1995) has demonstrated that this assumption does not imply any essential loss of generality. This means that any multiperiod random utility model can be approximated arbitrarily closely by random a utility model generated from max-stable utility processes.

The concept "continuity in probability" is a continuity property and it means that the probability that $\left|V_j(s) - V_j(t)\right| > \delta$, for any $\delta > 0$, tends towards zero as s tends towards t. In our context continuity in probability is a rather plausible property, and it does not necessarily imply that the samples paths of the process $\left\{V_j(t), t \geq a\right\}$ are continuous. For example, a stochastic process with jumps may still be continuous in probability. The continuity-in-probability condition only implies that the jumps of the process cannot occur "too frequently". The separability property is a very weak mathematical regularity condition. We refer to textbooks in probability theory for more details about separability. The independence condition stated in Axiom 1 is ad hoc from a theoretical perspective. However, since there are only two states the assumption that the processes V_1 and V_0 are independent does not seem restrictive, since only utility differences matters for the agent's choices.

Assumption 1 implies that the utility $V_j(t)$ has the structure

(2.1) $$V_j(t) = v_j(t) + \varepsilon_j(t)$$

where $v_j(t)$ is a deterministic term and $\varepsilon_j(t)$ is a random term with c.d.f. that is independent of $v_j(t)$ at time t.

[3] There seems to be some confusion in the statistical literature as regards notation. What some authors call type III extreme

Axiom 2

Provided there is no learning and structural state dependence (i.e., preferences are exogenous) the indirect utility $\max_k V_k(t)$ *is independent of the choice history* $\{J(s), s \le t\}$.

Axiom 2 states that in the "reference case" when there are no state dependence effects, the indirect utility at time *t* is uncorrelated with current and past choices. (Here it is of course understood that the structural processes $\{v_j(t), t \ge 0\}$, $j = 0,1$, are given). This is in fact a natural characterization of exogenous preferences. The intuition is as follows: Consider the indirect utility $\max_k V_k(t)$ and *J*(*t*) at time *t*. The indirect utility $\max_k V_k(t)$ is completely determined by preferences and choice constraints (prices, income and choice set). Knowledge about *J*(*t*) provides no *additional* information that should be of value for the level of $\max_k V_k(t)$, since $\max_k V_k(t)$ is the highest utility the agent can attain. If $\max_k V_k(t)$ were correlated with the choice *J*(*t*) this would mean that not all information that is relevant for the agent's indirect utility is captured in the utility function, and consequently the utility function would be ill defined. To further facilitate interpretation consider a large number of independent replications of a choice experiment. "Independent" means here that the respective random terms are drawn independently across experiments. In different replications the choices and the indirect utilities may be different due to different draws of the random terms. However, if the indirect utilities and choices in each experiment turn out to be correlated it would mean that the c.d.f. of the indirect utility, given that state 1 (say) was chosen would be different from the corresponding c.d.f. given that state 0 was chosen. This appears inconsistent with the fact that the agent's perception about the value of each alternative is fully represented by the indirect utility function. Thus if Axiom 2 is not fulfilled this means that the utilities contain random elements that do *not* represent individual randomness (in the sense of Thurstone), but are related to unobservables known to the agent while perceived as unobservables by the researcher. In Section 9 we shall discuss how one can allow for such (additional) random effects.

Before we introduce the next assumption we need an additional definition. By *stationary environment* we mean that the observed covariates that influence the utility processes are constant over time.

Axiom 3

---

value distributions, others call type I. Here we have adopted the convention of Resnick (1987).

In a stationary environment the utility processes are stationary when time is large.

This is an intuitively plausible assumption. The reason why we require time to be large is that there may be possible “upstarting” effects, which will gradually fade away.

Consider now the implications of the above assumptions. For simplicity we first discuss the discrete time case. When Axioms 1 to 3 hold one can demonstrate that (in discrete time) that the utilities $\{V_j(t), t > 0\}$ are updated according to

$$V_j(t) = \max\left(V_j\left(t-1\right) - \theta, W_j(t)\right) \tag{2.2}$$

where $W_j(t) = w_j(t) + \eta_j(t), w_j(t)$, is a deterministic term and $\eta_j(t),\ j = 0,1, t = a, a+1, a+2,...,$ are i.i.d. random variables that are independent of $\{w_j(t)\}$ with c.d.f.

$$P\left(\eta_j(t) \le x\right) = \exp\left(-\exp\left(-x\right)\right) \tag{2.3}$$

and θ is a non-negative parameter. The proof is given in the appendix. If we drop the requirement that the utility processes generated by (2.2) should tend to a stationary process when $W_j(t), t = a, a+1,...,$ then we may also have $\theta = 0$.

The variable $W_j(t)$ can be interpreted as the period *t*-specific utility of alternative *j*. The term θ may be interpreted as a *preference discount* factor. In the stationary case $w_j(t)$ does not depend on *t*. When θ is large there would be no dependence on the past and in this case $V_j(t) = W_j(t)$.

A stochastic process with the property (2.2), and with $\theta = 0$, is called an *extremal* process in probability theory, see for example Resnick (1987). We shall call the process that satisfies (2.2) with $\theta \ge 0$, a *modified extremal process*. Thus, the modification consists in allowing for preference depreciation represented by θ.

3. Properties and interpretation of extremal utility processes

Although our theoretical justification of the model rests entirely on the intuitive content of Axioms 1 to 3, it would nevertheless strengthen the behavioral rationale for the model if additional motivation and interpretation can be provided. This is the purpose of the present section.

Note first that when the utility process is a modified extremal process it is, possibly apart from a deterministic depreciation, a pure jump stochastic process. This means that the "current" utility has the role of an "anchoring" level such that unless the values of new stimuli exceed the anchoring level, utility will not be updated. The deterministic depreciation means that the anchoring effect at a given point in time gradually fades away as time passes. The interpretation above is consistent with results from psychological research where it is typically found that individuals do respond to stimuli only if they are sufficiently strong. Moreover, it is found that anchoring effects are rather common.

(a) Bounded rationality and sensory threshold

Above we have, except for the unemployment state, interpreted the jumps of the utility of being employed as due to concrete job arrivals. However, we may also give an alternative interpretation of the "jump" property of (2.2). In fact, representations like (2.2) have a long history in psychology and measurement theory and stem from empirical evidence indicating that individuals seem not to react to stimuli unless their intensity exceeds some threshold, see Gescheider (1997). In the present case this means that the agent will not pay attention to the "stimulus", represented by $W_j(t)$, unless $W_j(t)$ exceeds the "threshold", $V_j(t)-\theta$. Thus, with this interpretation, if

$$W_j(t) \leq V_j\left(t-1\right)-\theta \text{ ,}$$

the agent will not notice the new stimulus and will, accordingly, not update his preferences (apart from depreciation). This interpretation is similar to Fechner's (1860/1966) notion of "just noticeable differences" (jnd). The moments of time updating occurs may not necessarily relate to actual events, but could be due to sudden glimpses of insight at which epochs it is realized that utility re-evaluations are needed because the pleasure derived from the respective states is not the same as it used to be. This interpretation is perhaps of particular relevance for the unemployment state since the notion of non-market opportunities is more vague and abstract than the corresponding market ones. The special case with $\theta = 0$ corresponds to the situation in which the agent has "perfect" memory in the sense that previous preference evaluations are retained perfectly fixed in the agents mind. Thus, in this case no re-evaluation takes place unless the current stimulus associated with alternative j exceeds the previous utility value.

Recall that the jump property of the utility process is a derived property that essentially stems from Axiom 2. This indicates a curious link between our concept of exogenous preferences as characterized in Axiom 2 and the psychological literature, in which some evidence support the sensory threshold hypothesis mentioned above.

(b) Recall of job - and non-market opportunities

Here we assume that the agent faces a stream of arriving jobs- and non-market opportunities. In the discrete time case considered here only one job-and one non-market opportunity arrive in each period. As discussed above, the arrival of non-market opportunities may not necessarily correspond to actual concrete physical alternatives but rather to an updating of values according to psychological states of awareness, enabling the agent to broadening his horizon and taking into account new possibilities. We consider the case in which the agent has total recall of all jobs -as well as non-market opportunities that have arrived. For expository simplicity, we first consider the case with discrete time and with $\theta = 0$. Remember that in each period *t* the agent can move frictionless between the two states. Recall that $V_j(t)$ is the utility of being in state *j* in period *t*. Let $W_1(t)$ be the utility of the job offer that arrives in period *t*. Then, since there is no depreciation and the agent has recall, the value of the highest offer equals

(3.1) $$V_1(t) = \max\left(V_1\left(t-1\right), W_1(t)\right) = \max_{k \le t} W_1(k) \ .$$

In the case with depreciation the argument is completely analogous. Similarly, we view the search in the unemployment state as one where the agent updates his utility of being unemployed in each period. Let $W_0(t)$ denote the instantaneous utility of being unemployed in period t. When evaluating this utility the agent takes into account the distribution of the values of the incoming job offers, $\{W_1(k), k \ge t\}$, as well as the value of leisure (and not-market activities) enjoyed as unemployed. If there is no depreciation the value of being unemployed is, similarly to (3.1), equal to

(3.2) $$V_0(t) = \max\left(V_0\left(t-1\right), W_0(t)\right) = \max_{k \le t} W_0(k) \ .$$

This means that the agent's preferences are such that new offers that are not better than all previous rejected offers will never be accepted. Thus the case with no depreciation ($\theta = 0$) represents an extreme case where preferences are such that previous rejected opportunities ones and for all are classified as uninteresting by the agent. When depreciation is allowed (as in (2.2)), it is possible that the agent will accept offers that are worse than previously rejected offers. This happens in period t if $W_1(t) > W_0(t)$ and $W_1(t) > \max\left(V_1(t-1), V_0(t-1)\right) - \theta$.

(c) Job search without recall

We shall now show that (2.2) also has an intuitive interpretation in the case with no recall. Consider first the case with $\theta = 0$. If the agent is employed in state *j* at time $t-1$ he is free to accept the new job offer to state *j* at time *t*. As above, let $W_1(t)$ be the utility of the job offer that arrives at time *t*, and similarly $W_0(t)$ the utility of the non-market opportunity that arrives at time t. Remember that the agent is (by assumption) able to move frictionless between states. With no loss of generality suppose now that the agent is in state 1 at period $t-1$. Then his utility at time $t-1$ is $V_1(t-1)$. If $W_1(t) > V_1(t-1)$, then in this case the agent's highest utility of remaining in state 1 is $W_1(t)$. If, however, $W_1(t) \le V_1(t-1)$ the agent's utility of being in state 1 is $V_1(t-1)$. We can write this more compactly as

$$V_1(t) = \max\left(V_1(t-1), W_1(t)\right). \tag{3.9}$$

The utility of state 0 at period $t-1$ is $V_0(t-1)$. Suppose the arriving non-market opportunity at period t has utility $W_0(t)$ such that $W_0(t) \le V_0(t-1)$. Due to (3.3) we have that $V_1(t) \ge V_1(t-1)$, and due to the assumption that at time $t-1$ the agent was in state 1, this implies that $V_1(t) \ge W_0(t)$, and consequently no transition will occur. Thus, we can define $V_0(t) = V_0(t-1)$ when $W_0(t) \le V_0(t-1)$, since this yields $V_1(t) > V_0(t)$, which is consistent with no transition. Consider next the case when $W_0(t) > V_0(t-1)$. Then the agent will go to state 0 provided $W_0(t) > V_1(t)$, otherwise he will reject the new non-market offer and remain in state 1. We therefore can define $V_0(t) = W_0(t)$ in this case, because this is consistent with the fact that $V_0(t) > V_1(t)$. This means that we can, similarly to the case for state 1, represent the utility of being in state 0 as

$$V_0(t) = \max\left(V_0(t-1), W_0(t)\right). \tag{3.4}$$

The argument is completely similar when $\theta > 0$. We have thus shown that (2.2) has an intuitive interpretation also in the case with no recall. The continuous time version is completely analogous.

(d) Interpretation under de Haan's max-spectral representation in continuous time

Due to de Haan (1984), the property that the indirect utility processes are extremal (or modified extremal processes) in continuous time allows for a particular interpretation that we shall discuss in this section. Let T_{1k} denote the arrival time of job offer k, where $k = 1, 2, \ldots,$ represents an arbitrary enumeration (not necessarily ordered). Let U_{1k} denote the utility of job offer k. Define

$$U_{1k} = u_1(T_{1k}) + \xi_{1k}. \tag{3.5}$$

Suppose now that the pair $(T_{1k}, \xi_{1k}), k = 1, 2, \ldots,$ are the points of a bivariate Poisson process with intensity

$$\lambda(t)e^{-\varepsilon}.$$

This means that the points are independently scattered and the probability that a point *k* with $T_{1k} \in (t, t+dt), \xi_{1k} \in (\varepsilon, \varepsilon + d\varepsilon)$ equals

$$\lambda(t)e^{-\varepsilon}dt\,d\varepsilon.$$

If we define $\tilde{W}_1(s,t)$ and $V_1(t)$ by

(3.6) $$\tilde{W}_1(s,t) = \max_{s<T_{1k}\le t}\left(U_{1k} - (t - T_{1k})\theta\right)$$

and

(3.7) $$V_1(t) = \max_{a\le T_{1k}\le t}\left(U_{1k} - (t - T_{1k})\theta\right)$$

we realize that one can update $\{V_1(t)\}$ as

(3.8) $$V_1(t) = \max\left(V_1(s) - (t-s)\theta, \tilde{W}_1(s,t)\right).$$

A completely similar relation holds for the utility of the unemployment state. One can easily show that $\tilde{W}_1(s,t)$ is type III extreme value distributed. Moreover, since the Poisson points are independently scattered it follows immediately that $V_1(t)$ and $\tilde{W}_1(s,t)$ are independent. From de Haan (see Resnick, 1987) it follows that $\{V_1(t), t \ge a\}$ is a modified extremal process. The alternative representation of the modified extremal utility process introduced here means that one can interpret $\tilde{W}_1(s,t)$ as the utility of the most attractive discounted job offer that arrives within the interval (*s*,*t*). In this set up the probability that more than one job offer with arrival time within $(t, t+dt)$, with associated error term within $(\varepsilon, \varepsilon + d\varepsilon)$, is negligible. However, in contrast to the usual one- dimensional Poisson process, jobs with other associated error terms will arrive within the infinitesimal time interval $(t, t+dt)$. Moreover, Dagsvik (2006) demonstrates that we have

(3.9) $$E\tilde{W}_1(s,t) = \log\left[\int_s^t \exp\left(w_1(\tau) - (t-\tau)\theta\right)d\tau\right]$$

where

(3.10) $$w_1(t) = u_1(t) + \log \lambda(t) .$$

It follows from (3.9) that for small Δt

$$\log E\tilde{W}_1\left(t-\Delta t, t\right) \cong w_1(t) + \log(\Delta t)$$

which means that $w_1(t)$ can (apart from an additive constant) be interpreted as the mean utility of most attractive job offer that arrives within $\left(t-\Delta t, t\right)$.

4. Implications for the choice probabilities in continuous time when transitions out of employment are voluntary

We shall now explore the implications from the theory above for the structure of the choice probabilities in the case where transitions out of employment are voluntary. The more general and realistic case where an employed worker may lose his job will be considered in Section 6.

If the utility processes are independent and modified extremal processes as represented in (3.8) Dagsvik (1988) has proved that the choice process $\{J(t), t \geq a\}$ is a Markov chain in continuous time. (This holds in the general case with non-stationarity). See also Resnick and Roy (1990) who have generalized the results of Dagsvik (1988). For an overview of the properties of Markov chains we refer to Feller (1968). Recall that a denotes the "starting point in time" (age), i.e., the first time epoch the agent is exposed to labor market choices. Dagsvik (1988, 2002) show that the corresponding transition probabilities $\{Q_{ij}(s,t)\}$ can be expressed as

(4.1) $$Q_{ij}\left(s,t\right) = P\left(J(t)=j \,\middle|\, J(s)=i\right) = \frac{\int_s^t \exp\left(w_j(\tau) - \left(t-\tau\right)\theta\right) d\tau}{\sum_{k=0}^{1} \int_a^t \exp\left(w_k(\tau) - \left(t-\tau\right)\theta\right) d\tau}$$

for $a \leq s \leq t, i, j = 0,1, i \neq j$, and $Q_{ii}\left(s,t\right) = 1 - Q_{ij}\left(s,t\right)$. Furthermore, the probability of being in state j is given by

(4.2) $$P_j(t) = P\left(J(t)=j\right) = \frac{\int_a^t \exp\left(w_j(\tau) - \left(t-\tau\right)\theta\right) d\tau}{\sum_{k=0}^{1} \int_a^t \exp\left(w_k(\tau) - \left(t-\tau\right)\theta\right) d\tau}.$$

One can also show that $v_j(t),\ j = 0,1$, given in (2.1) can be expressed as

$$v_j(t) = \log\left(\int_a^t \exp\left(w_j(\tau) - (t-\tau)\theta \right) d\tau \right).$$

Thus, in the two-state case the Axioms above provide a theoretical basis for assuming the Markov property. This seems to be plausible: When there is no state dependence there is no reason why the transition probabilities (or equivalently the transition intensities) should depend on past choice history.

Next we shall consider the corresponding transition intensities, or hazard functions. Recall that the hazard functions of a continuous time Markov chain are defined (usually) as

$$q_{ij}(t) = \lim_{s \to t} \frac{Q_{ij}(s,t)}{t-s},$$

for $i \neq j$. Let $T_i(s)$ be the duration of stay in state i, given that the agent entered state i at time s. It now follows immediately from (4.1) and (4.2) that the corresponding transition intensities of the Markov chain $\{J(t), t \geq a\}$ are given by

(4.3)
$$q_{ij}(t) = \frac{\exp\left(w_j(t) \right)}{\sum_{k=0}^{1} \int_a^t \exp\left(w_k(\tau) - (t-\tau)\theta \right) d\tau},$$

for $i \neq j$. Furthermore, the c.d.f. of $T_i(s)$ is given by

(4.4)
$$P\left(T_i(s) > y \right) = \exp\left(-\int_s^{s+y} q_{ij}(\tau) d\tau \right).$$

The parameter θ is closely linked to the serial dependence of the utility processes, cf. Dagsvik (2002). Let $V(t) = \max_k V_k(t)$ be the indirect utility. One can show that

(4.5)
$$Corr\left(V(s), V(t) \right) = \zeta\left(\frac{\sum_k \int_a^s \exp\left(w_k(\tau) + \theta\tau \right) d\tau}{\sum_k \int_a^t \exp\left(w_k(\tau) + \theta\tau \right) d\tau} \right)$$

where the function $\zeta(x), x \in [0,1]$, equals

(4.6) $$\zeta(x) = -\frac{6}{\pi^2}\int_0^x \frac{\log z}{1-z}\,dz\,,$$

see Tiago de Oliveira (1973) and Dagsvik (1988). It can be shown that this function is positive, continuous and strictly increasing on [0,1] with $\zeta(0)=0$ and $\zeta(1)=1$.

Consider the case with time independent systematic utility component $w_j(t)$, so that $w_j(t)=w_j$. Then, (4.2) reduces to

(4.7) $$P_j(t) = P_j = \frac{\exp\left(w_j\right)}{\sum_{k=0}^{1}\exp\left(w_k\right)}$$

and (4.1) reduces to

(4.8) $$Q_{ij}\left(s,t\right) = P_j \cdot \frac{1-\exp\left(-\left(t-s\right)\theta\right)}{1-\exp(-(t-a)\theta)}$$

for $s \le t$, $i \ne j$, and

(4.9) $$Q_{ii}(s,t) = \frac{\exp\left(-\left(t-s\right)\theta\right)-\exp\left(-(t-a)\theta\right)}{1-\exp\left(-(t-a)\theta\right)} + P_i \cdot \frac{1-\exp\left(-\left(t-s\right)\theta\right)}{1-\exp\left(-(t-a)\theta\right)}\,.$$

The autocorrelation function (4.5) in this case reduces to

(4.10) $$Corr\left(V(s),V(t)\right) = \zeta\left(1-\frac{1-\exp\left(-\left(t-s\right)\theta\right)}{1-\exp\left(-(t-a)\theta\right)}\right).$$

The corresponding transition intensities are given by

(4.11) $$q_{ij}(t) = \frac{\theta P_j}{1-\exp(-(t-a)\theta)}$$

for $i \ne j$. From (4.8) to (4.11) we see that the transition probabilities and intensities become stationary when $t\theta$ is large. However, when $(t-a)\theta$ is small this is not so. In fact, the transition probabilities and intensities increase by the factor $1\big/\left(1-\exp\left(-(t-a)\theta\right)\right)$ when $(t-a)\theta$ is small. The interpretation is that when the agent is very "young" the choice history is very short and therefore the effect of taste persistence is weak. As the agent grows older this "upstarting" effect disappears gradually, and becomes negligible when $\exp\left(-(t-a)\theta\right)$ is close to zero. This is also seen from the autocorrelation function in (4.10). When $(t-a)\theta$ is large, the autocorrelation function is, apart from a strictly increasing transformation, equal to $\exp\left(-\left(t-s\right)\theta\right)$. However, when $(\mathrm{t}-\mathrm{a})\theta$ is "small" the autocorrelation is influenced by the term $1/\left(1-\exp\left(-t\theta\right)\right)$ in such a way that it becomes weaker when t increases while $t-s$ is kept constant.

Consider the special case when θ tends towards zero. Then, by using $\mathrm{l'H\hat{o}pital's\ rule}$ (4.8) and (4.9) become equal to

(4.12) $$Q_{ij}(s,t)=\frac{t-s}{t-a}\cdot P_j\,,$$

for $i\neq j$, and

(4.13) $$Q_{ii}(s,t)=\frac{s-a}{t-a}+\frac{t-s}{t-a}\cdot P_i.$$

In this case the autocorrelation function becomes,

(4.13) $$Corr\left(V(s),V(t)\right)=\zeta\left(\frac{s-a}{t-a}\right)=\zeta\left(1-\frac{(t-s)}{t-a}\right).$$

Thus, when time is large there are no transitions in this case. However, when *t* is "small" the autocorrelation will be less than 1, which means that when the agent is "young" the effect of taste persistence is reduced because the choice history is short and the autocorrelation is therefore less than one.

In the special case where $w_j(t)$ does not depend on *t* and *t* is large the transition probabilities and intensities reduce to

(4.15) $$Q_{ij}\left(s,t\right)=P_j\left(1-\exp\left(-\left(t-s\right)\theta\right)\right)$$

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

(4.16) $$Q_{ii}\left(s,t\right)=\exp\left(-\left(t-s\right)\theta\right)+P_i\left(1-\exp\left(-\left(t-s\right)\theta\right)\right)$$

where P_j is given in (4.7). The corresponding transition intensities are given by

(4.17) $$q_{ij}(t)=\theta P_j$$

for $i\neq j$.

The depreciation effect represented by the parameter θ can in fact be given an interesting alternative interpretation, which we shall now point out. Specifically, one can interpret the depreciation mechanism as a stochastic device where the habit persistence effect is represented by means of a particular Poisson process. The intuition is that, at independent random points Z_k in time, $k = 1, 2, ...,$ (random to the observer), the agent forgets or stops caring about previous evaluations and only takes into account new stimuli. That is, if for some k, $Z_{k-1} < t < Z_k$, the agent will, at time t, only take into account previous preference evaluations within the interval $\left(Z_{k-1}, t\right]$. The intensity of this process is θ, which means that the probability that for some k, $Z_k \in (t, t+dt)$, equals θdt. If the agent *does* care, he will keep the previous value *without* depreciation. Consider now (4.15). Since under stationarity no explanatory variable change over time, the agent will thus never change to a new alternative unless he stops taking the past into account. Specifically, to change state from time s to time t the agent must stop taking into account previous values of the alternatives at some point in time within the interval (s, t). Since the interval between two events in a Poisson process is exponentially distributed with θ as parameter, the probability that this will happen is $1 - \exp(-(t-s)\theta)$. The conditional probability that the agent shall choose alternative j, given that he stops caring about the past, equals P_j, since the corresponding utilities that govern this choice are $W_j(t) = w_j + \eta_j(t),\ j = 0,1$, where $\eta_j(t),\ j = 0,1$, are independent extreme value distributed as in (2.3). Hence, by multiplying P_j by the probability that the agent will begin to neglect the past some time within (s, t), (4.15) is obtained. Similarly, if the agent occupies state i at time s he will continue to be in state i at time t if he cares about past preference evaluations (which has probability $\exp\left(-\theta(t-s)\right)$), or if he does not care about past evaluations (with probability $1 - \exp\left(-\theta(t-s)\right)$) but chooses state i with probability P_i. Hence, eq. (4.16) follows. Eq. (4.17) states that the intensity of transition from i to j can happen when the agent stops taking the past into account in $\left(t, t+dt\right)$, which happens with probability θdt. Given that he forgets about the past he will go to state j with probability P_j. Hence, we obtain that the probability of going from state i to j in $\left(t, t+dt\right)$ equals $\theta dt P_j$.

5. Using a particular invariance postulate to characterize the utility of unemployment

So far our theory is silent on how the utility of being unemployed is determined. In general, the application of search theory to this end in empirical research is controversial because economic theory is silent about functional form and besides, there is no general agreement among researchers about the correct theory. In addition, if one believes that bounded rationality is important, it is not clear how one should apply rational search theoretic arguments in this case. Consequently, in contrast to search theory as a point of departure, we shall in this section discuss an alternative approach. This approach is analogous to recent advances in theoretical psychophysics. In psychophysics there is a tradition, starting with Luce (1959), which attempts to characterize functional form of mathematical representations by imposing specific invariance properties (see also Falmagne, 1985). Dagsvik and Strøm (2005), and Dagsvik et al. (2006), have demonstrated how such principles can be successfully applied. We shall now proceed by making a particular invariance postulate. Before we state the next assumption we first need to introduce additional notation. We shall now write the respective instantaneous utility and utility of being unemployed as $W_0(t,\lambda)$ and $V_0(t,\lambda)$ to emphasis that these terms may depend on the job offer arrival intensity λ. For simplicity, we only consider the discrete time case.

Axiom 4

Let λ_1, λ_2, $\lambda_1^{\bullet}$ and $\lambda_2^{\bullet}$ be job arrival intensities such that

$$(5.1)\qquad P\Big(W_0\big(t,\lambda_2\big)<V_0\big(t-1,\lambda_1\big)-\theta\Big)<P\Big(W_0\big(t,\lambda_2^{\bullet}\big)<V_0\big(t-1,\lambda_1^{\bullet}\big)-\theta\Big)$$

for $s<t$, and suppose that $w_0(\tau,\lambda)$, for given λ, is independent of τ for $\tau\le t$. Then for all $\tau>0$,

$$(5.2)\qquad P\Big(W_0\big(t,\tau\lambda_2\big)<V_0\big(t-1,\tau\lambda_1\big)-\theta\Big)<P\Big(W_0\big(t,\tau\lambda_2^{\bullet}\big)<V_0\big(t-1,\tau\lambda_1^{\bullet}\big)-\theta\Big).$$

The empirical counterpart of the left hand side of (5.1) is the fraction of persons that evaluate the instantaneous utility of being unemployed in period *t* as lower than the depreciated utility of being unemployed in period $t-1$ when the arrival rate of job offers is constant and equal to λ_1 up to time $t-1$ and changes to λ_2 from period $t-1$ to *t*, and the right hand side is similar with λ_1 and λ_2 replaced by $\lambda_1^{\bullet}$ and $\lambda_2^{\bullet}$. The left hand side of (5.1) is equivalent to the fraction of persons that do not change their utility (apart from depreciation) of being unemployed when the job offer arrival intensity changes from λ_1 to λ_2, and the right hand side is analogous. Eq. (5.1) states that when the job arrival intensity changes from λ_1 to λ_2 the corresponding fraction of agents that do not change their utility (apart from depreciation) of being unemployed is less than the fraction of agents that do not change their utility of being unemployed when the job arrival intensity changes from $\lambda_1^{\bullet}$ to $\lambda_2^{\bullet}$. Thus, Axiom 4 states that if the fraction of agents that do not change their utility (apart from depreciation) of being unemployed when the arrival intensity changes from λ_1 to λ_2 is less then the fraction of agents that do not change their utility (apart from depreciation) of being unemployed when the arrival rate change from $\lambda_1^{\bullet}$ to $\lambda_2^{\bullet}$, this inequality remains true when all the job arrival intensities are rescaled by a positive factor. In other words, it is assumed that as long as the *relative changes* remain constant the aggregate behavior of the workers will not change in a "fundamental way". That is, although the respective fractions of peoples that change their employment decisions may change, this change will not reverse the inequalities in Axiom 4.

In the appendix we prove that Axiom 4 implies that the deterministic component, $w_0(t,\lambda)$ of $W_0(t,\lambda)$ takes the functional form (as function of λ)

(5.3) $$w_0(t,\lambda) = b(t) + \rho\left(\lambda^{\kappa} - 1\right)\big/\kappa \,,$$

where $b(t)$ is a suitable function of t that does not depend on λ, and $(\lambda^{\kappa}-1)/\kappa = \log\kappa$ when $\kappa = 0$. Also κ and ρ may depend on t, but we shall in the following assume that they are constants. However, not all parameter values for ρ and κ are acceptable. When λ tends towards infinity it is reasonable to require that $w_0(t,\lambda)$ tends towards a limit equal to the value of employment because when λ is large search costs will be negligible. On the other hand, when λ tends towards zero one expects that $w_0(t,\lambda)$ should tend towards infinity. As a result (5.3) should have the form

(5.4) $$w_0(t,\lambda) = \bar{u}_1(t) + \rho\lambda^{\kappa} \,,$$

where ρ and κ are negative and $\bar{u}_1(t)$ is the expected value of employment.

Next we shall see that (5.4) in fact is consistent with a typically search theoretic formulation. Consider an unemployed agent. He is viewed as being uncertain about his opportunities in the labor market and about the utility of arriving job offers. Job offers arrive according to a Poisson process with arrival intensity λ (possibly individual specific) assumed known by the individual. The corresponding search cost per unit of time is denoted c. Let w_0 denote the utility of searching and assume that the agent is boundedly rational in the sense that she ignores discounting and is unable to account for the possibility of lay-off. Let U_1 be the value of an arriving job offer and let $w_1 = EU_1$. Then, applying the standard Bellman type of argument, cf. Lippman and McCall (1981), or Burdett et al. (1984), we get

(5.5) $$w_0 = \left(1-\lambda\Delta t\right)\left(w_0 - c\Delta t\right) + \lambda\Delta t\left(E\max\left(U_1, w_0\right) - c\Delta t\right) + o(\Delta t) \,.$$

Equation (5.5) says that (in a stationary environment) the utility of searching for work is evaluated as follows. When searching two things can happen in a small time interval of length Δt. With probability $1-\lambda\Delta t$ no job offer arrives within $(t, t+\Delta t)$ so that the expected utility in this case remains equal to w_0 minus search cost $c\Delta t$. Otherwise, a job offer arrives with probability $\lambda\Delta t$, in which case expected utility equals $E\max(w_0, U_1) - c\Delta t$. After re-arranging, dividing by Δt and letting Δt tend towards zero we obtain

(5.6)
$$w_0 = E\max(w_0, U_1) - \frac{c}{\lambda}.$$

Note furthermore that we can write

(5.7)
$$\begin{aligned} &E\max(w_0, U_1) = w_0 P(U_1 < w_0) + E(U_1 | U_1 > w_0) P(U_1 > w_0) \\ &= w_0 (1 - P(U_1 > w_0)) + E(U_1 | U_1 > w_0) P(U_1 > w_0). \end{aligned}$$

When inserting (5.7) into (5.6) it follows that we can re-write (5.6) as

(5.8)
$$w_0 = \bar{w}_1 - \frac{c}{\lambda^*},$$

where $\lambda^* = \lambda P(U_1 > w_0)$ and $\bar{w}_1 = E(U_1 | U_1 > w_0)$. Eq. (5.8) states that the utility of searching is equal to the expected value of working, given an acceptable job, minus the expected cost until an acceptable job arrives. With suitable interpretation of the job arrival rate we realize that (5.8) is consistent with (5.4) with $\kappa = -1$. When $P(U_1 > w_0) = 1$, (5.8) reduces to

(5.9)
$$w_0 = EU_1 - \frac{c}{\lambda} = w_1 - \frac{c}{\lambda},$$

which is consistent with (5.4) without reinterpretation of the job arrival rate. This special case corresponds to a setting in which the agent will accept the first offer arriving. This corresponds to the actual policy at many public unemployment agencies, where the employment manager selects jobs viewed as “suitable” for the respective unemployed worker. Normally, the unemployed worker cannot refuse a job offer viewed as suitable by the manager, if he wishes to continue to receive unemployment benefits.

Note that the conventional search theoretic formulation above rests on the assumption that utility is additively separable. This is a very strong assumption that requires justification. However, such justification is seldom provided. Second, the assumption that agents know the arrival rate- and the distribution of the utilities of the job offers, and are able to make the type of calculations reviewed above, is very strong and controversial.

6. An extension that allows for involuntary transitions from employed to unemployed

In this section we shall consider the case where the agent may lose his job when employed. We shall discuss how the framework above can be modified to cover this case. When employed we shall still assume that the arrival rate of market opportunities does not depend on whether or not the agent is employed.

Axiom 5

The agents may lose their jobs according to a Poisson process with intensity $\mu(t)$. *That is,* $\mu(t)$ *is the intensity of involuntary transitions into unemployment. The arrival intensity of job offers does not depend on the duration of employment or whether the agent is unemployed or employed.*

Axiom 5 simply states that the chance of losing the job does not depend on tenure, but is allowed to depend on time through business cycle indicators. Moreover, the arrival rate of job offers is assumed to be independent of whether or not the agent is employed. From Axioms 1 to 5 we prove in the appendix that the transition probabilities in this case become

$$(6.1) \qquad Q_{01}(s,t) = \frac{\exp\left(-M(t)\right)\int_s^t \exp\left(w_1(x) + x\theta + M(x)\right)dx}{\int_a^t \left[\exp\left(w_0(x) + x\theta\right) + \exp\left(w_1(x) + x\theta\right)\right]dx}$$

and

$$(6.2) \qquad Q_{10}(s,t) = 1 - Q_{01}(s,t) - \frac{\exp\left(M(s) - M(t)\right)\int_a^s \left[\exp\left(w_1(x) + x\theta\right) + \exp\left(w_0(x) + x\theta\right)\right]dx}{\int_a^t \left[\exp\left(w_0(x) + x\theta\right) + \exp\left(w_1(x) + x\theta\right)\right]dx}$$

where

$$M(t) = \int_a^t \mu(x)dx \, .$$

Note that $\exp(-M(t))$ is the probability that the worker, if employed, will not lose his job involuntarily before time *t*. The formulae for the transition probabilities given in (6.1) and (6.2) are useful in situations where the researcher only has fragmentary data on the states the agents occupy at some discrete points in time, because they imply that one can estimate the underlying parameters without knowing the whole (constrained) choice history of the agents. Similarly, they can be used to make predictions (policy simulations) given the parameters of the model.

7. Identification

An interesting question is whether or not the structural parameters θ, $w_0(t)$ and $w_1(t)$ are identified from the transition intensities. In the case with involuntary transitions, let $r_{01}(t)$ and $r_{10}(t)$ denote the transition intensities out of unemployment and out of employment, respectively. We have that $r_{01}(t) = q_{01}(t)$ and $r_{10}(t) = q_{10}(t) + \mu(t)$. In the appendix we show that the terms $w_0(t) + t\theta$ and $w_1(t) + t\theta$ can be recovered through the following relations

(7.1) $$w_1(t) + \theta t = \log r_{01}(t) + \int_d^t \left(r_{01}(x) + r_{10}(x)\right) dx - \int_d^t \mu(x) dx + \kappa$$

and

(7.2) $$w_0(t) + \theta t = \log(r_{10}(t) - \mu(t)) + \int_d^t \left(r_{01}(x) + r_{10}(x)\right) dx - \int_d^t \mu(x) dx + \kappa$$

where $d \geq a$ and κ are arbitrary constants. To identify θ one needs to assume that there is a time period, say from t_1 to t_2, where $w_1(t)$ is constant for $t \in [t_1, t_2]$ (or alternatively $w_0(t)$ is constant in some interval). If this is the case θ can be determined by the equation

(7.3) $$(t_2 - t_1)\theta = \log r_{01}(t_2) - \log r_{01}(t_1) + \int_{t_1}^{t_2} \left(r_{01}(x) + r_{10}(x)\right) dx - \int_{t_1}^{t_2} \mu(x) dx \, .$$

In the case with only two choice alternatives, an immediate consequence of (7.1) and (7.2) is that our theory imply no restrictions on the choice process $\{J(t), t \geq a\}$ beyond the Markov property. This is also true when $\mu(t) = 0$. However, the particular reparametrization and corresponding interpretation of the transition intensities of this Markov chain in terms of θ , $w_1(t)$ and $w_0(t)$ allows the researcher to make causal inferences and to assess the impact of related policy interventions. If $\mu(t)$ is not known one cannot obtain nonparametric identification as above.

Consider the issue of identification under (3.6) and (5.4). Let $y_1(t)$ denote the right hand side of (7.1) and $y_2(t)$ the right hand side of (7.2), which we assume are known. Let $m(t) = \log\lambda(t)$ Then, by (3.6), (5.4), (7.1) and (7.2) we get

(7.4) $$m(t) = y_1(t) - u_1(t) - \theta t$$

and

(7.5) $$\bar{u}_1(t) + \rho \exp\left(\kappa m(t)\right) + \theta t = y_2(t) .$$

Now suppose for example that $u_1(t)$ is proportional to the conditional mean of the logarithm of the wage rate, given schooling and experience and assume furthermore that $\bar{u}_1(t)$ is parametrized as a function of known time dependent covariates. Since $m(t)$ is not observed we insert (7.4) into (7.5). Then, with sufficient variation in $y_1(t)$ and $y_2(t)$ and the other covariates one can determine κ, ρ, θ, $\bar{u}_1(t)$ and the proportional factor in $u_1(t)$.

8. Structural state dependence

Dagsvik (1998, 2002) discuss how a framework such as the one discussed above can be modified to allow for state dependence in the sense that choices and/or involuntary transitions affect the structural part of the utility function. In this case the structural part of the utility function becomes endogenous and it is not obvious how the corresponding choice probabilities shall be calculated in this case. For simplicity, let time be discrete and suppose now that $w_j\left(t, h(t)\right)$ depends in some way on the choice history $h(t)$ up to time $t-1$, while $\eta_j(t),\ j = 0,1,$ are independent of $h(t)$. Since $w_1\left(t, h(t)\right)$ consists of two components the dependence could enter through $u_1(t)$ or through $\lambda(t)$, or both, and similarly for $w_0\left(t, h(t)\right)$. In this case Dagsvik (1998) demonstrates that the transition intensities can be obtained simply by substituting the structural terms $\left\{w_j(t)\right\}$ in the transition intensities given above, by the corresponding state dependent terms, $\left\{w_j\left(t, h(t)\right)\right\}$. Unfortunately, the nice formulae for the transition probabilities, analogous to (6.1) and (6.2), will no longer hold in this case since obviously the choice process $\left\{J(t), t > 0\right\}$ will no longer be Markovian.

The significance of providing a framework that allows the researcher to identify structural state dependence effects may be realized by the following example: In the literature on the analysis of the determinants of unemployment researchers have argued that short term economic policies that alleviate unemployment tend to lower aggregate unemployment rates in the long run by preventing the loss of work-enhancing labor market experience (cf. Phelps, 1972, for example). This argument rests on the assumption that unemployment has a real and lasting effect on the future probability of unemployment of the currently unemployed. In contrast, other authors (see for example Cripps and Tarling, 1974), have advanced the opposite view by assuming that agents differ in their propensity to experience unemployment and in their unemployment duration times, and that these differences cannot be fully accounted for by measured variables. In addition, they maintain that the actual experience of having been unemployed or the duration of past unemployment does not affect future incidences or duration. As a result, short-term economic policies will have no effect on long-term unemployment in their model. As is well known, one cannot in general distinguish between structural state dependence and heterogeneity in preferences and choice constraints by using statistical methods alone. Additional theoretical assumptions are needed. See also Heckman (1981a, 1981b, 1991). The framework proposed in this paper enables the researcher to test the two competing theories provided the unknown parameters of the model are the same for all individuals in the sample. This means that unobserved heterogeneity is completely captured by the error terms of the utility function represented by the modified extremal processes (which includes the randomness associated with the Poisson arrival process of job and non-market opportunities), with structural terms that only vary across the population by observed characteristics. In addition, the random Poisson process of involuntary entrance to unemployment governed by $\mu(t)$ also only varies across the population by observed characteristics. The reason why it is possible to identify effects due to structural state dependence is because our theory above postulates how the hazard rates look like in the reference case with no state dependence effects. Dagsvik (1998, 2006) discusses details on this identification issue.

9. Random effects (unobserved heterogeneity)

The stochastic model developed above is assumed to rationalize randomness at the individual level. Recall that by this we understand randomness due to (i) opportunities that arrive at random points in time with uncertain values (as perceived by the agent), (ii) updating of information about the distribution of uncertain events and variables, and (iii) the taste-shifters of the utility function may vary randomly due to the agent having difficulties with assessing definite values to the respective opportunities once and for all.

To the observing econometrician additional randomness may occur because some variables known to the individual are not observed by the researcher. In this section we discuss briefly the case where selected unknown parameters of the model are allowed to be individual specific (random effects). In the proportional hazard modeling framework a positive random term that is supposed to represent unobserved heterogeneity is typically introduced as a multiplicative component of the hazard rate out of unemployment. In the framework above, the representation of unobserved heterogeneity is not as straight forward as in the proportional hazard rate model. This is due to the fact that there are several parameters to whom (additive) random effects can be assigned. Furthermore, the issue on how these random effects should be distributed is a delicate one. One solution is to use non-parametric discrete distributions for the random effects. Similarly to Heckman and Singer (1984), Gaure, Røed and Zhang (2005), to mention just a few, one can approximate the true mixing distribution with a non-parametric discrete distribution. Jain et al. (1994) have demonstrated that this can be done with several random effects. To be specific, let $\boldsymbol{b} = \left(b_1, b_2, ..\right)$ denote a vector consisting of m (say) random effects and let $\Omega = \left\{\boldsymbol{b}^1, \boldsymbol{b}^2, ..., \boldsymbol{b}^K\right\}$ denote the set of support vector points in R^m. Let g_k be the probability mass at point $\boldsymbol{b}^k$. The estimation problem we face is to determine the number of support points K, their locations $\left\{\boldsymbol{b}^k\right\}$ and the mixing probabilities g_k. Note that in this setup all the m random effects are allowed to be interdependent in a completely general way. Let $L(\boldsymbol{b})$ denote the conditional likelihood function given the random effect vector $\boldsymbol{b}$. Then the unconditional log likelihood function is given by

$$\log\left(\sum_{k=1}^{K} L(\boldsymbol{b}^k) g_k\right).$$

Lindsay (1995) has treated inference issues in this type of nonparametric mixing models.

An alternative and rather interesting specification of random effects in the general non-stationary case is obtained by modifying the terms $w_j(t),\ j = 0,1$, to contain a particular additive random effect generated from a particular Stable distribution. Recall that a stable distribution is characterized by four parameters, namely $\alpha \in (0,1]$, $\beta \in [-1,1]$, $\sigma > 0$ and $\gamma \in (-\infty, \infty)$. The parameter α characterizes the fatness of the tail(s) in the sense that decreasing values imply increasing fatness, β represents the skewness where $\beta = 0$ yields a symmetric c.d.f. while $\beta = 1(-1)$ yields a distribution that is totally skew to the right (left). The parameter σ is a scale parameter that is similar to the standard deviation. However, σ^2 can only be interpreted as the variance in the special case where $\alpha = 2$, in which case the stable distribution reduces to the Normal distribution. If $\alpha < 2$, the variance does not exist. The parameter γ is a location parameter that is equal to the expectation if $\alpha > 1$. If $\alpha \le 1$ the expectation does not exist. When $\alpha < 1$, $\beta = 1$ and $\gamma \ge 0$ the corresponding stable random variable takes values on the positive part of the real line. A notation that is often used to represent a stable distribution is $S_\alpha(\sigma,\beta,\gamma)$. Specifically, let θ be distributed according to $S_\alpha\left(b^{1/\alpha},1,0\right)$, where b is a positive constant and $\alpha < 1$. (For information about stable distributions, see Samorodnitsky and Taqqu, 1994). To avoid complication in the presentation we only consider the discrete time case. The continuous time case is however completely analogous. The modification of the structural terms, $\{w_j(t)\}$, is done by replacing $w_j(t)$ by $\tilde{w}_j(t) = w_j(t) + \log Z_j(t)$, where $Z_j(t), t = 1,2,..., j = 1,2,$ and have c.d.f. $S_\alpha(\sigma,1,0)$ with $\alpha < 1$, which means that $Z_j(t)$ are positive. Conditional on the random effects the hazard function out of employment (in discrete time) becomes, according to (4.3) equal to

$$\tilde{q}_{ij}(t) = \frac{\exp\left(\tilde{w}_j(t)\right)}{\sum_{k=0}^{1}\sum_{\tau=a}^{t}\exp\left(\tilde{w}_k(\tau) - (t-\tau)\theta\right)d\tau}, \tag{9.2}$$

for $i \ne j$. Consequently, the c.d.f. of the duration of employment becomes

$$P\left(T_i(s) > y\right) = E\exp\left(-\sum_{\tau=s}^{y}\tilde{q}_{ij}(\tau)d\tau\right). \tag{9.3}$$

Now, one can show that the choice process under this type of random effect is still Markovian, and that (see Dagsvik, 2006)

$$E\exp\left(-\sum_{\tau=s}^{y}\tilde{q}_{ij}(\tau)d\tau\right) = \exp\left(-\sum_{\tau=s}^{y}q^{\bullet}_{ij}(\tau)d\tau\right) \tag{9.4}$$

where

$$q^{\bullet}_{ij}(t) = \frac{\exp\left(\alpha w_j(t)\right)}{\sum_{k=0}^{1}\sum_{\tau=a}^{t}\exp\left(\alpha w_k(\tau) - (t-\tau)\alpha\theta\right)d\tau}. \tag{9.5}$$

We notice the remarkable property that the functional form in (9.5) is the same as the one in (4.3) (modified to the discrete time case), with the instantaneous utilities and the depreciation parameter re-scaled by the parameter α, which we recall is less than one. As a result the new autocorrelation parameter $\alpha\theta$ is less than the corresponding depreciation parameter θ in the model without random effect. In other words, the autocorrelation in the indirect utility function increases when this type of random effect is introduced in the model. However, the Markovian property still holds in this case. This also means that the autocorrelation parameter θ can represent the effect of a particular form of random effect. This is so because when $\theta \geq 5$, (say), autocorrelation will be negligible whereas if α is small $\alpha\theta$ will be small which implies that the autocorrelation may be substantial.

10. Two examples of empirical model specification

Example 1: Only data on single spell duration of unemployment available

In this section we discuss briefly the issue of empirical modelling in the special case in which the researcher only has information about duration in the unemployment state. Thus, we shall only consider the modelling of the hazard function for transitions out of the unemployment state. Recall that the instantaneous utility of working, $W_1(t)$, can be interpreted as the utility of the highest job offer that arrives in the period. Hence, it is a function of the wage rate among other things. We therefore assume that it depends linearly on an instrument equation for the wage rate. This instrument equation can be estimated on the basis of data that contains information about wages. Suppose now that $w_j(t),\ j=0,1$, change slowly over time (relative to tθ). By first order Taylor expansion around t we have that

$$\text{(10.1)} \qquad \log\left(\exp\left(w_0(\tau)\right)+\exp\left(w_1(\tau)\right)\right)=\log\left(\exp\left(w_0(t)\right)+\exp\left(w_1(t)\right)\right)+(\tau-t)\zeta(t)$$

where

$$\text{(10.2)} \qquad \zeta(t)=\frac{\exp\left(w_0(t)\right)w_0'(t)+\exp\left(w_1(t)\right)w_1'(t)}{\exp\left(w_0(t)\right)+\exp\left(w_1(t)\right)}.$$

Consequently, under this approximation it follows from (4.3) and (10.2) that we can write the hazard function for transitions out of unemployment as

$$\text{(10.3)} \qquad \begin{aligned} q_{01}(t) &= \frac{\exp\left(w_1(t)\right)}{\sum_{k=0}^{1}\int_a^t \exp\left(w_k(\tau)-\left(t-\tau\right)\theta\right)d\tau} \cong \frac{\exp\left(w_1(t)\right)}{\sum_{k=0}^{1}\exp(w_k(t)-\theta t-t\zeta(t))\int_a^t \exp\left(\tau(\theta+\zeta(t))\right)d\tau} \\ &= \frac{(\theta+\zeta(t))\exp(w_1(t))}{[1-\exp\{-(t-a)(\theta+\zeta(t))\}]\sum_{k=0}^{1}\exp(w_k(t))} \cong \frac{(\theta+\zeta(t))\exp(w_1(t))}{\sum_{k=0}^{1}\exp(w_k(t))}. \end{aligned}$$

From the last expression we realize that if $\zeta(t)$ varies slowly over time it will simply be absorbed in the parameter θ.

Consider next the extension where the utility of unemployment is allowed to depend on the remaining time the worker is eligible for unemployment benefits. This implies that the utility of unemployment depends on duration, *d* (say). When duration dependence is introduced this represents a form of structural state dependence. According to the discussion in section 8 we can still apply the formulas for the hazard function (4.3) with $w_0(t)$ replaced by the corresponding duration dependent version $w_0(t,d)$. If Axiom 4 holds we can write the systematic part of the instantaneous utility of being unemployed as

(10.4) $$w_0(t,d) = \bar{u}_1(t,d) + \rho\lambda(t)^{\kappa}.$$

Since we only have information about durations in the unemployment state we cannot identify and estimate $\lambda(t)$ non-parametrically. When we take into account (3.10), (10.4) and the duration dependency we can express (10.3) as

(10.5) $$q_{01}(t,d) = \frac{\left(\theta + \zeta(t,d)\right)\exp\left(u_1(t)\right)}{\exp\left(u_1(t)\right) + \exp\left(\bar{u}_1(t,d) + g(t)\right)}$$

where

(10.6) $$g(t) = \rho\lambda^{\kappa} - \log\lambda(t)$$

and $\zeta(t,d)$ is obtained by replacing $w_0(t)$ and $w_0'(t)$ in (10.2) by $w_0(t,d)$ and $\partial w_0(t,d)/\partial t$, respectively. If $\zeta(t,d)$ is small compared to θ we can simply substitute $\zeta(t,d)+\theta$ by θ, which implies that

(10.7) $$q_{01}(t,d) = \frac{\theta\exp\left(u_1(t)\right)}{\exp\left(u_1(t)\right) + \exp\left(\bar{u}_1(t,d) + g(t)\right)}.$$

One way of checking if θ is much larger than $\zeta(t,d)$ is to estimate the model based on the specification (10.7) and then afterwards compute $\zeta(t,d)$ by means of (10.2). In case $\zeta(t,d)$ is not small compared to θ one may try to use the exact representation of the model, that is, a specification based on (4.3). This may however be cumbersome, because one cannot express the model solely through $w_0(t,d) - w_1(t)$. This implies that one needs to allow for a representation of $\log\lambda(t)$ in addition to $g(t)$ in the model. This may entail a difficult identification problem.

An alternative approach in this case is to estimate the model in two stages: In the first stage estimate the model based on (10.7) and subsequently compute $\zeta(t,d)$ based on (10.2). In the second stage estimate the model based on (10.5) with the estimate of $\zeta(t,d)$ as an additional explanatory variable.

Example 2: Data available on both employment and unemployment durations for some individuals
In this example the researcher is assumed to have a sample of observations where durations on employment and unemployment spells are observed for all, or some, individuals. For simplicity, assume that the individuals enter the workforce as unemployed, and after some time enter the employment state where they remain until they become unemployed and remain unemployed until the sampling period is finished, or they remain employed until the sampling period is finished. More general data sets can be analysed in a completely similar way as the one we shall outline below. We assume here that the researcher is ignorant about whether or not exit from employment to unemployment is voluntary of involuntary. Thus, the transition intensity out of employment is in this case equal to $q_{10}(t)+\mu(t)$. Let t_{i1} denote the point in time when individual i obtains employment the first time and t_{i2} the point in time when individual i exits from employment. Let t_3 denote the right sensoring time epoch. Let S_1 be the subsample of individuals with $t_{i2}<t_3$, where transitions out of employment are voluntary and let S_2 be the subsample of individuals with $t_{i2}<t_3$, where transitions out of employment are involuntary. The likelihood for the individuals in S_1, L_1 (say), is equal to

(10.8)
$$\begin{aligned} L_1 &= \prod_{i\in S} \exp\left(-\int_a^{t_{i1}} q_{01}(\tau)d\tau\right) q_{01}\left(t_{i1}\right) \exp\left(-\int_{t_{i1}}^{t_{i2}} \left(q_{10}(\tau)+\mu(\tau)\right)d\tau\right) q_{10}\left(t_{i2}\right) \exp\left(-\int_{t_{i2}}^{t_3} q_{01}(\tau)d\tau\right) \\ &= \prod_{i\in S} q_{01}\left(t_{i1}\right) q_{10}\left(t_{i2}\right) \exp\left(-\int_a^{t_{i1}} q_{01}(\tau)d\tau - \int_{t_{i2}}^{t_3} q_{01}(\tau)d\tau - \int_{t_{i1}}^{t_{i2}} \left(q_{10}(\tau)+\mu(\tau)\right)d\tau\right). \end{aligned}$$

Similarly, the likelihood for those individual who do not belong to S_2, L_2, equals

(10.9)
$$L_2 = \prod_{i\in S} q_{01}\left(t_{i1}\right) \mu\left(t_{i2}\right) \exp\left(-\int_a^{t_{i1}} q_{01}(\tau)d\tau - \int_{t_{i2}}^{t_3} q_{01}(\tau)d\tau - \int_{t_{i1}}^{t_{i2}} \left(q_{10}(\tau)+\mu(\tau)\right)d\tau\right).$$

11. Conclusion

In this paper we have discussed the issue of functional form in structural models for discrete panel data analysis with emphasis on duration of stay and transitions between employment and unemployment. We have demonstrated that one can use a particular axiomatic approach to motivate the choice of functional form for the choice model that accounts for randomness at the individual level. The most important property implied by the axioms is that the choice model (choice of state) is a Markov chain in continuous time. Furthermore, our theoretical set up produces a behavioral decomposition and interpretation of the transition intensities (hazard function). In the two-state case we consider here it is important to notice that our theory implies no additional restrictions on the hazard functions in the sense that there is a one to one correspondence between any set of hazard functions and the underlying structural parameters. We have also proposed a particular invariance assumption to obtain a characterization of the utility of unemployment as a function of the arrival intensity of job offers. We have furthermore considered extensions that allow for structural state dependence, involuntary transitions and random effects.

Appendix

Proof of eq. (6.1) and (6.2):

Recall that the agents are assumed to lose their jobs according to a Poisson process. Recall also that the choice process $\{J(t), t>0\}$ (with voluntary transitions) is a Markov chain. As a result, the modified choice process that allows for involuntary transitions is also a Markov chain. Consider an agent that is in state 0 at time $t+\Delta t$ given that he was in state *i* at time $s<t$. This can occur if the agent was in state 0 at time *t* and no transition occurred in $(t, t+\Delta t)$, which has probability $1-\Delta t q_{01}(t)+o(\Delta t)$, or if the agent was in state 1 at time *t* and moved to state 0 in $(t, t+\Delta t)$. The latter event has probability $\Delta t\left(q_{10}(t)+\mu(t)\right)+o(\Delta t)$. Formally, this can be expressed as

(A.1)
$$\begin{aligned} Q_{i0}\left(s, t+\Delta t\right) &= \left(1-\Delta t q_{01}(t)+o(\Delta t)\right) Q_{i0}(s,t)+\left(\Delta t\left(q_{10}(t)+\mu(t)\right)+o(\Delta t)\right) Q_{i1}(s,t) \\ &= \left(1-\Delta t q_{01}(t)\right) Q_{i0}(s,t)+\Delta t\left(q_{10}(t)+\mu(t)\right) Q_{i1}(s,t)+o(\Delta t). \end{aligned}$$

After reorganizing terms and inserting $Q_{i1}(s,t)=1-Q_{i0}(s,t)$, we obtain that

(A.2)
$$\frac{Q_{i0}\left(s, t+\Delta t\right)-Q_{i0}\left(s,t\right)}{\Delta t}+\left(q_{01}(t)+q_{10}(t)+\mu(t)\right) Q_{i0}(s,t)=q_{10}(t)+\mu(t)+\frac{o(\Delta t)}{\Delta t}.$$

By passing to the limit; $\Delta t \to 0$ we obtain

(A.3)
$$\frac{\partial Q_{i0}\left(s,t\right)}{\partial t}+\left(q_{01}(t)+q_{10}(t)+\mu(t)\right) Q_{i0}(s,t)=q_{10}(t)+\mu(t).$$

By a completely similar argument we obtain that

(A.4)
$$\frac{\partial Q_{i1}\left(s,t\right)}{\partial t}+\left(q_{01}(t)+q_{10}(t)+\mu(t)\right) Q_{i1}(s,t)=q_{01}(t).$$

From (A.3), (A.4) and (4.3), (6.1) and (6.2) follow by standard methods for solving differential equations, see for example Berck and Sydsæter (1993).

Q.E.D.

Proof of eq. (7.1) and (7.2):

From (4.3) it follows that

(A.5)
$$\int_d^t \left(r_{10}(x)+r_{01}(x)\right)dx = \int_d^t \frac{\sum_{j=0}^{1} \exp\left(w_j(x)+x\theta\right)dx}{\sum_{j=0}^{1}\int_a^x \exp\left(w_j(\tau)+\tau\theta\right)d\tau} + \int_d^t \mu(x)dx$$
$$= \Big|_d^t \log\left(\sum_{j=0}^{1}\int_a^x \exp\left(w_j(\tau)+\tau\theta\right)d\tau\right) = \log\left(\sum_{j=0}^{1}\int_a^t \exp\left(w_j(\tau)+\tau\theta\right)d\tau\right) + \int_d^t \mu(x)dx - \kappa$$

where κ is a constant. Hence we obtain that

(A.6) $$\log r_{01}(t) + \int_a^t \left(r_{01}(x)+r_{10}(x)\right)dx + \kappa = \log\left(r_{01}(t)\sum_{j=0}^{1}\int_a^t \exp\left(w_j(\tau)+\tau\theta\right)d\tau\right) - \int_d^t \mu(x)dx = w_1(t)+t\theta$$

which proves (7.1). The proof of (7.2) is completely similar.

Q.E.D.

Proof of eq. (5.3):

It follows readily from (2.2) that when $w_0(\tau,\lambda_1)$ is constant up to period $t-1$, then

(A.7) $$\tilde{P}(\lambda_1,\lambda_2) \equiv P\left(W_0(t,\lambda_2) < V_0(t-1,\lambda_1)-\theta\right) = \frac{K(t)\exp\left(w_0(\lambda_1)\right)}{K(t)\exp\left(w_0(\lambda_1)\right)+\exp\left(w_0(\lambda_2)\right)}$$

where

$$K(t) = \sum_{k=a}^{t-1} \exp\left(-(t-k)\theta\right).$$

Hence we can express (A.7) as

(A.8) $$\tilde{P}(\lambda_1,\lambda_2) = \tilde{F}\left(\psi(\lambda_1)/\psi(\lambda_2)\right)$$

where $\psi(\lambda) = \exp\left(w_0(\lambda)\right)$ and

$$\tilde{F}(x) = \frac{xK(t)}{xK(t)+1}.$$

Note that $\tilde{F}$ is continuous and strictly increasing. By applying Theorem 14.19 in Falmagne (1985), p. 338, we realize that Axiom 4 implies that

(A.9) $$\tilde{P}(\lambda_1,\lambda_2) \equiv \tilde{F}\left(h(\lambda_1)/h(\lambda_2)\right) = F\left(\frac{\delta_1\left(\lambda_1^{\alpha}-1\right)-\delta_2\left(\lambda_2^{\alpha}-1\right)}{\alpha}\right)$$

where $\delta_j > 0,\ j = 1,2,$, and α are constants and *F* is a continuous and strictly increasing function. From (A.9) it follows that

(A.10) $$w_0(\lambda_1) - w_0(\lambda_2) = G\left(\frac{\delta_1\left(\lambda^{\alpha}-1\right)-\delta_2\left(\lambda^{\alpha}-1\right)}{\alpha}\right)$$

where *G* is a continuous function that is given by $G(x) = \log\left(\tilde{F}^{-1}F(x)\right)$. By letting $\lambda_1 = \lambda_2 = \lambda$ in (A.10) we get

(A.11) $$G\left((\delta_1-\delta_2)\left(\lambda^{\alpha}-1\right)/\alpha\right) = 0$$

for all $\lambda > 0.$ Since G is not constant for all arguments, (A.11) implies that $\delta_1 = \delta_2 = \delta$ (say). Let

$$x = \frac{\delta\left(\lambda_1^{\alpha}-1\right)-\delta\left(\lambda_2^{\alpha}-1\right)}{\alpha} \quad \text{and} \quad z = \frac{\delta\left(\gamma^{\alpha}-1\right)-\delta\left(\lambda_1^{\alpha}-1\right)}{\alpha}$$

where $\gamma > 0$ is fixed. From (A.10) we obtain that

$$w_0(\gamma) - w_0(\lambda_1) = G(z) \quad \text{and} \quad w_0(\gamma) - w_0(\lambda_2) = G(x+z),$$

from which it follows that

(A.12) $$G(z) + G(x) = G(x+z).$$

Eq. (A.12) is a Cauchy type of functional equation whose only continuous solution is $G(x) = cx$, where *c* is a constant, see for example Falmagne (1985), Theorem 3.2. As a result we get from (A.10) with $\lambda_1 = \lambda$, that

(A.13) $$w_0(\lambda) = \frac{\beta\left(\lambda^{\alpha}-1\right)}{\alpha} + w_0(1).$$

This completes the proof.

Q.E.D.

Proof of eqs. (2.2) and (3.8):

We shall use results obtained in Dagsvik, Jia and Strøm (2006) to prove (2.2). To this end, note that Axioms 1 and 2 are completely equivalent to Axioms 1, 2 and Condition 1 in Dagsvik, Jia and Strøm (2006). The only difference is that in the present paper the time parameter is replaced by income as parameter, and the type III extreme value marginal distributions are replaced by type I extreme value distributions (Fréchet distributions) in Dagsvik, Jia and Strøm (2006).

It is readily verified that if a random variable X is Fréchet distributed then $\log X$ is type III extreme value distributed. Now assume that the requirement of type III marginal c.d.f. in Axiom 1 is replaced by Fréchet marginal c.d.f. Let $\{V_j^*(t), t > a\}$ be the corresponding utility processes. Then we realize from the proof of Lemma 2 and Theorem 1 in Dagsvik, Jia and Strøm (2006), that we can write

(A.15) $$V_j^*(t) = \max\left(V_j^*(t-1)b_t, W_j^*(t)\right)$$

where b_t is a positive term and $W_j^*(t), t = a+1, a+2, ...,$ are independent and Fréchet distributed. Apart from the factor b_t, (A.15) is equivalent to (2) in Dagsvik, Jia and Strøm (2006). The role of Axiom 3 in Dagsvik, Jia and Strøm (2006) is to ensure that $b_t = 1$. Let $V_j(t) = \log V_j^*(t)$. Then from (A.15) we get that

(A.16) $$V_j(t) = \max\left(V_j(t-1) - \theta_t, W_j(t)\right)$$

where $\theta_t = -\log b_t$ and $W_j(t) = \log W_j^*(t)$. Since $W_j^*(t)$ is Fréchet distributed it follows that $W_j(t)$ is type III extreme value distributed and can accordingly be written as $W_j(t) = w_j(t) + \eta_j(t)$ where $\eta_j(t), t = a+1, a+2, ...,$ are i.i.d. with c.d.f. $\exp\left(-e^{-x}\right)$ for real x, and $w_j(t)$ is a deterministic term.

Now is follows from (A.16) that

(A.17)
$$\begin{aligned}
&P\left(V_j(t-1) \le u_1, V_j(t) \le u_2\right) = P\left(V_j(t-1) \le u_1, V_j(t-1) \le u_2 + \theta_t, W_j(t) \le u_2\right) \\
&= P\left(V_j(t-1) \le \min\left(u_1, u_2 + \theta_t\right)\right) P\left(W_j(t) \le u_2\right) \\
&= \exp\left(-\exp\left(v_j(t-1) - \min\left(u_1, u_2 + \theta_t\right)\right) - \exp\left(w_j(t) - u_2\right)\right) \\
&= \begin{cases} \exp\left(-\exp\left(v_j(t-1) - u_1\right) - \exp\left(w_j(t) - u_2\right)\right) & \text{for} \quad u_1 < u_2 + \theta_t, \\ \exp\left(-\left[\exp\left(v_j(t-1) - \theta_t\right) + \exp\left(w_j(t)\right)\right]\exp\left(-u_2\right)\right) & \text{for} \quad u_1 \ge u_2 + \theta_t. \end{cases}
\end{aligned}$$

Consider the joint c.d.f. (A.17) for $u_1 < u_2 + b_t$. With no loss of generality assume now that the environment is stationary and that time is large. Due to Assumption 3, $v_j(t-1)$ and $w_j(t)$ must be independent of time. But then it also follows from the case where $u_1 \geq u_2 + \theta_t$ that also $\theta_t = \theta$ must be constant. From (A.16) we get, under Axiom 3, that

(A.18) $$V_j(t) = \max\left(\eta_j(t), \eta_j(t-1) - \theta, \eta_j(t-2) - 2\theta, ..., \eta_j(a+1) - (t-a-1)\theta\right) + w_j$$

which implies that

(A.19) $$P\left(V_j(t) \leq u\right) = \prod_{k=0}^{t-a-1} \exp\left(-\exp\left(w_j - u - k\theta\right)\right) = \exp\left(-e^{w_j - u}\left[1 - e^{-(t-a)\theta}\right] \Big/ \left(1 - e^{-\theta}\right)\right).$$

Eq. (A.19) shows that stationarity is achieved when $t-a$ is large and $\theta > 0$. It also shows that stationarity cannot be achieved unless $\theta > 0$. This completes the proof for the discrete time case. In the continuous time case Tiago de Oliveira (1973) has demonstrated that the representation (3.8) implies a strictly stationary (Markovian) stochastic process.

Q.E.D.

References

Berck, P. and K. Sydsæter (1993): *Economists' Mathematical Manual*. Springer-Verlag, Berlin.

Burdett, K., N. Kiefer, D.T. Mortensen, and G. Neumann (1984): Earnings, Unemployment and the Allocation of Time over Time. *Review of Economic Studies*, **51**, 559-578.

Cripps, T. F. and R. J. Tarling (1974): An Analysis of the Duration of Male Unemployment in Great Britain 1932-1973. *The Economic Journal*, **84**, 289-316.

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. (1995): How Large is the Class of Generalized Extreme Value Random Utility Models? *Journal of Mathematical Psychology*, **39**, 90-98.

Dagsvik, J. K. (1998): On the Structure of Behavioral Multistate Duration Models. Memorandum 98/11, Department of Economics, University of Oslo.

Dagsvik, J. K. (2002): Discrete Choice in Continuous Time: Implications of an Intertemporal Version of the IIA Property. *Econometrica*, **70**, 817-831.

Dagsvik, J. K. (2006): Functional Form of Hazard rated in Behavioral Multistate Duration Models. Mimeo, (work in progress). Statistics Norway.

Dagsvik, J. K., and S. Strøm (2005): Sectoral Labor Supply, Choice Restrictions and Functional Form. To appear in *Journal of Applied Econometrics.*

Dagsvik, J. K., S. Strøm and Z. Jia (200): Utility of Income as a Random Function: Behavioral Characterization and Empirical Evidence. *Mathematical Social Science,* **51**, 23-57.

Debreu, G. (1958): Stochastic Choice and Cardinal Utility. *Econometrica*, **26**, 440-444.

Feller, W. (1968): *An Introduction to Probability Theory and Its Applications*, Vol. I. Wiley, New York.

de Haan, L. (1984): A Spectral Representation for Max-Stable Processes. *Annals of Probability*, **12**, 1194-1204.

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

Fechner, G. T. (1860/1966): *Elemente der Psychophysik*. Druck und Verlag von Breitkopfs Hartel, Leipzig. [*Elements of Psychophysics*, (Vol I). Holt, Rinehart & Winston, 1966, New York].

Feller, W. (1968): *An introduction to Probability Theory and Its Applications,* Vol. I. Wiley, New York.

Gescheider, G. A. (1997): *Psychophysics: The Fundamentals.* Lawrence Erlbaum Associates, Inc., Publishers, New Jersey.

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, USA.

Heckman, J.J. (1981b): Heterogeneity and State Dependence. In S. Rosen (ed.), *Studies in Labor Markets*. University of Chicago Press, Chicago.

Heckman, J. J., and B. Singer (1984): Econometric Duration Analysis. *Journal of Econometrics*, **24**, 63-132.

Heckman, J.J. (1991): Identifying the Hand of the Past: Distinguishing State Dependence from Heterogeneity. *American Economic Review*, **81**, 75-79.

Jain, D. C., N. J. Vilcassim and P. K. Chintagunta (1994): A Random-coefficient Logit Brand-Choice Model Applied to Panel Data. *Journal of Business and Economic Statistics*, **12**, 317-328.

Lindsay, B. G. (1995): *Mixture Models: Theory, Geometry and Applications.* Regional Conference Series in Probability and Statistics. Institute of Mathematical Statistics, Hayward, CA, USA.

Lippman, S. A., and J. J. McCall (1981): The Economics of Uncertainty: Selected Topics and Probabilistic Methods. In K. J. Arrow and M. D. Intriligator (eds.), *Handbook of Mathematical Economics*, Vol. I, Ch. 6, North-Holland, Amsterdam.

Luce, R. D. (1959): On the Possible Psychophysical Laws. *The Psychological Review*, **66**, 81-95.

Luce, R. D. (1996): The Ongoing Dialog between Empirical Science and Measurement Theory. *Journal of Mathematical Psychology*, **40**, 78-98.

Phelps, E. (1972): *Inflation Policy and Unemployment Theory.* Norton, New York.

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

Resnick, S. I., and R. Roy (1990): Multivariate Extremal Processes, Leader Processes and Dynamic Choice. *Advances in Applied Probability*, **22**, 309-331.

Gaure, S. K. Røed, K., and T. Zhang (2005): Time and Causality: A Monte Carlo Assessment of the Timing -of-Events Approach. Mimeo, The Ragnar Frisch Centre for Economic Research, Oslo.

Thurstone, L. L. (1927): A Law of Comparative Judgment. *Psychological Review*, **34**, 273-286.

Tiago de Oliveira, J. (1973): An Extreme Markovian Stationary Process. *Proceedings of the Fourth Conference in Probability Theory* (Brasov, 1971), pp. 217-225. Editura Acad. R. S. R. Bucharest (1973), Romania.

Tversky, A. (1969): Intransitivity of Preferences. *Psychological Review*, **76**, 31-48.