Dagsvik, J. (2002): " Discrete choice in continuous time : implications of an intertemporal version of the IIA property". Econometrica, Vol. 70, No. 2 (March, 2002), 817–831

DOI: http://dx.doi.org/10.1111/1468-0262.00307

Title:	Discrete choice in continuous time : implications of an intertemporal version of the IIA property
Author:	Dagsvik, John K.
Version:	"The copyright to this article is held by the Econometric Society, http://www.econometricsociety.org/. It may be downloaded, printed and reproduced only for personal or classroom use. Absolutely no downloading or copying may be done for, or on behalf of, any for-profit commercial firm or other commercial purpose without the explicit permission of the Econometric Society. For this purpose, contact Claire Sashi, General Manager, at sashi@econometricsociety.org.
Publisher:	The Econometric Society (Copyright holder) and Wiley (commercial publisher)
Source:	Econometrica

DISCRETE CHOICE IN CONTINUOUS TIME: IMPLICATIONS OF AN INTERTEMPORAL VERSION OF THE IIA PROPERTY

BY JOHN K. DAGSVIK[1]

This paper proposes a particular behavioral axiom to characterize the stochastic structure of static discrete choice models with serially correlated utilities. This assumption extends Luce's axiom; "Independence from Irrelevant Alternatives", to the intertemporal setting. Under general regularity conditions the implication of this assumption is that the individual choice process is a Markov chain with transition probabilities that have a particularly simple structure. It is also discussed how the framework can be extended to deal with structural state dependence.

KEYWORDS: Choice over time, state dependence, Markovian choice processes, extremal processes, random utility processes, independence from irrelevant alternatives.

1. INTRODUCTION

During the last decades significant progress has been made in developing and applying (static) discrete choice models. Some of these analyses were based on earlier theoretical contributions by Luce (1959), and Luce and Suppes (1965). Specifically, Luce (1959) proposed the axiom known as "Independence from Irrelevant Alternatives" (IIA). Although the IIA assumption is strong, and is known not to hold in many choice settings, it nevertheless represents an intuitive and powerful principle of stochastic rationality.[2]

In this paper we propose an intertemporal version of the IIA Axiom (IIIA) and we derive the implications for the corresponding choice model in continuous time. The point of departure is a static- or myopic random utility setting where utilities are serially correlated due to temporally persistent unobservables. It is, even in a static setting, of interest to obtain a characterization of serial dependence in preferences. First, if panel data is available one needs to account for serial correlation in the estimation procedure. Second, it may be of interest to predict the effect of policy simulations on transition rates from one alternative to another. This axiom (IIIA) can be described as follows: Consider the particular case in which there are no structural state dependence as a result of previous choice experience and the past choice sets are all the same but where the choice set in the current period is expanded to include new alternatives that were never feasible before. Then IIIA states that choices among the new alternatives that enter the choice set are independent of any choice in the past. The intuition is that even if previous choices provide information about preferences over the alternatives in the "old" choice set, these choices provide no information about the utilities of the "new" alternatives, since they were not feasible in the past. Under suitable regularity conditions we demonstrate that IIIA implies a random utility representation where the utilities associated with each alternative are independent extremal processes, cf. Resnick (1987).

Now, it follows from Dagsvik (1983 and 1988) that when the choice sets are constant or increasing over time, the extremal utility processes yield a choice model which is a Markov chain (in continuous time), and where the transition probabilities have a particular structure as a function of the choice set and the parameters of the utility processes. Note that the markovian property does not follow immediately from IIIA, since IIIA is silent about situations where current choices also were feasible in the past.

The paper is organized as follows: In Section 2 the choice setting is formally described, the main assumptions are introduced and the implications for the preferences and the choice process derived. In Section 3 the implications for the choice probabilities are discussed. In Section 4 we consider an extension of the framework that allows for state dependence.

2. CHARACTERIZATION OF THE PREFERENCES AND THE INDIVIDUAL CHOICE PROCESS

In this section we discuss the basic assumptions and derive their implications for the random utility process. Specifically, we propose a behavioral assumption that enables us to characterize preferences in the case when there are no effects from past experiences on future preferences nor on future choice opportunities.

The individual decision-maker (agent) is supposed to have preferences over a finite set of alternatives. Future preferences are assumed random (to the agent himself) in the sense that they vary from one moment in time to the next in a way that cannot fully be predicted by the agent. Alternatively, one may interpret the utilities as deterministic to the agent but random to the observer due to variables that are perfectly foreseeable to the agent but unobserved by the analyst.

Let S be the index set of m alternatives, a_1, a_2,..., a_m, and let $\Im$ be the index set that corresponds to the collection of all non-empty subsets from S. We assume that S contains at least three alternatives. To each alternative, a_j, there is associated a stochastic process, $\{U_j(t), t \geq 0\}$, where $U_j(t)$ is the agent's (conditional indirect) utility of a_j given the information and choice history at time t. Moreover, $U_j(t) = v_j(t) + \varepsilon_j(t)$, where $v_j(t)$ is a deterministic component that may depend on alternative-specific attributes, and $\varepsilon_j(t)$ is a stochastic term. The agent chooses a_j at age t if $U_j(t)$ is the highest utility at t. Here age (time) is continuous. Let $\{J(t, B(t))\}$ denote the *choice process*, i.e.,

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

where $\left\{B(t), t > 0, B(t) \in \Im\right\}$ denotes the *choice set process*. If $B(t) = B(s)$ for all s and t, the choice set process is constant. Let $U(t) = \left(U_1(t), U_2(t), ..., U_m(t)\right)$, $\varepsilon(t) = \left(\varepsilon_1(t), \varepsilon_2(t), ..., \varepsilon_m(t)\right)$ and $v(t) = \left(v_1(t), v_2(t), ..., v_m(t)\right)$. We assume that $\left\{U(t)\right\}$ is separable and continuous in probability.[3] Moreover, we assume that the cumulative distribution function (c.d.f.) of $U(t)$ is absolutely continuous for any $t > 0$. This implies that there are no ties, that is

$$P\left(U_i(t) = U_j(t)\right) = 0.$$

When the finite dimensional distributions of the utility process $\left\{U(t), t > 0\right\}$ have been specified it is in principle possible to derive joint choice probabilities for a sequence of choices. However, since the class of intertemporal random utility models is quite large it is desirable to restrict this class on the basis of behavioral arguments. A related problem is that it seems to be rather difficult to find stochastic processes that are convenient candidates for utility processes in the sense that they imply tractable expressions for the choice probabilities in the intertemporal context.

One way of introducing structural restrictions into the model is to apply probabilistic versions of the assumption of rational behavior. A famous example of this type of assumption is Luce Choice Axiom; “Independence from irrelevant alternatives”, (IIA) (cf. Luce, 1959). A first attempt to extend IIA to the intertemporal setting was made by Dagsvik (1983).[4] Next we shall discuss the implications from another version of IIA, which is stated below.

AXIOM A1 (IIIA): *Let* $B(s) = B_1$, *for all* $s < t$, $B_1 \in \Im$, *and let* $B_2 \in \Im$ *be such that* $B_2 \setminus B_1 \neq \emptyset$.[5] *Then for* $j \in B_2 \setminus B_1$,

(2.1) $$P\left(J\left(t,B_2\right) = j \,\middle|\, J\left(s,B_1\right), \forall s < t\right) = P\left(J\left(t,B_2\right) = j\right).$$

It is important to stress that (2.1) does *not* mean that $\{J(t, B(t)), t > 0\}$ is a Markov chain, nor is it a Bernoulli process.[6] This is so because (2.1) is assumed to hold only when $j \notin B_1$ and is silent about the relationship between the choices at different points in time when $j \in B_1$. Axiom A1 states that when the most attractive element of B_2 is contained in $B_2 \setminus B_1$, the event, "a_j is the preferred alternative in B_2", is stochastically independent of the preference orderings in B_1 at time s, for $s < t$. It is therefore natural to interpret Axiom A1 as an intertemporal extension of the IIA property. The intuition is that even if previous choices provide information about the preferences over the alternatives in the "old" choice set, these choices provide no information about the utilities of the "new" alternatives, since they were not feasible in the past.

AXIOM A2: *At each point in time the distribution of the random vector,* $\varepsilon(t)$ *), does not depend on* $v(t)$.

AXIOM A3: *Let* $Z_j(t)$ *denote a vector of attributes specific to alternative j at time t that is differentiable with respect to t. The structural term* $v_j(t)$ *is given by* $v_j(t) = \tilde{v}\left(Z_j(t)\right)$ *where* $\tilde{v}(\cdot)$ *is a differentiable function that is known apart from a set of parameters and* *defined on a suitable set K. For any* $t > 0$*, and any real number x, there exists a value of* $Z \in K$ *such that* $\tilde{v}(Z) = x$.

Axiom A2 states that at each moment in time the random term of the utility function is independent of the structural term. Axiom A3 states that the structural term of the utility function can vary over the whole real line when attributes vary within *K*. Moreover, the attributes vary smoothly over time.

PROPOSITION 1: *Assume that A2 and A3 hold. Then for any* $B \in \Im$, *Axiom A1 implies that*

(2.2)
$$P\left(J\left(t,B\right)=j\right)=\frac{e^{\alpha v_j(t)}}{\sum_{k\in B} e^{\alpha v_k(t)}}$$

where $\alpha > 0$ *is an arbitrary constant.*

PROOF: Recall that $\left\{U(t), t > 0\right\}$ is continuous in probability. Recall also that since the utilities are independent of the choice set process, we are allowed to specify any sequence of choice sets which is useful for deriving implications about the preferences. To this end, let $B(t) = B$ and $B(t-) = B \setminus \{j\}$. Then A1 implies that

(2.3)
$$P\left(U_j(t)=\max_{k\in B} U_k(t), U_i(t-)=\max_{k\in B\setminus\{j\}} U_k(t-)\right)=P\left(U_j(t)=\max_{k\in B} U_k(t)\right)P\left(U_i(t)=\max_{k\in B\setminus\{j\}} U_k(t)\right).$$

By Theorem 50, p. 354, in Luce and Suppes (1965), (2.3) implies that the choice probabilities are given by the Luce model. Finally, Strauss (1979), pp. 42-43, has demonstrated that the parameters of the Luce model are related to the systematic part of the utility function as specified in (2.2), apart from an additive constant. *Q.E.D.*

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

Let us now proceed by investigating the intertemporal structure of the random utilities that follows from A1.

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

PROPOSITION 2: *Assume that A1 to A3 hold. If the utility processes,* $\{U_j(t), t \geq 0\}$, $j = 1, 2, ..., m$, *are independent at each point in time they have type III extreme value distributed marginals.*[7]

PROOF: If Theorem 5 of Yellott (1977) is combined with Proposition 1 the result of Proposition 2 follows. *Q.E.D.*

REMARK 2: Note that although there exists random utility models that satisfy IIA with interdependent utilities, there is, under IIA, no loss of generality in assuming that utilities are independent across alternatives.

AXIOM A4: *The utility processes* $\{U_j(t), t \geq 0\}$, $j = 1, 2, ..., m$, *are stochastically independent.*

REMARK 3: Recall that two stochastic processes $\{U_i(t), t \geq 0\}$ and $\{U_j(t), t \geq 0\}$ may be stochastically dependent even if $U_i(t)$ and $U_j(t)$ are stochastically independent at each point in time. For example, $U_i(t)$ and $U_j(s)$ may be dependent for $s \neq t$, even if $U_i(t)$ and $U_j(t)$ are independent. However, it seems plausible that in many applications the correlation between $U_i(t)$ and $U_j(s)$ is less than the correlation between $U_i(t)$ and $U_j(t)$, which implies that the utility processes are independent when the utilities at each point in time are independent.

AXIOM A5: *The utility process* $\{U_j(t), t \geq 0\}$, $j = 1, 2, ..., m$, *is a max-stable process.*

The next result is the main result of this paper.

Axiom A5 represents no loss of generality because Dagsvik (1995) has proved (under the absence of state dependence) that the choice probabilities of a multiperiod random utility model can be approximated arbitrarily closely by choice probabilities of the random utility model with max-stable utility processes. Recall that the max-stable processes have finite-dimensional marginal distributions that are multivariate extreme value distributions, cf. de Haan (1984).

THEOREM 1: *Assume A1 to A5. Then the utilities are extremal processes with type III extreme value marginal distribution.*

The proof of Theorem 1 is given in the appendix.

A description of the class of extremal processes can be found in Resnick (1987). An extension to inhomogeneous extremal processes has been made by Weissman (1975). For our purpose it will be convenient to work with a modified inhomogeneous extremal process. The modified extremal process differs from the (standard) inhomogeneous extremal process by a deterministic time trend. More precisely, a modified inhomogeneous extremal utility function will be defined as processes $\{U_j(t), t > 0\}$, $j = 1, 2, ..., m,$ given by

(2.4) $$U_j(t) = \max\left(U_j(s) - (t-s)\theta, W_j(s,t)\right),$$

$s < t,$ where $U_j(0) = -\infty,$ $\theta > 0$ is a constant and where $W_j(s,t)$ is independent of $U_j(s)$ and has cumulative distribution function

(2.5) $$P\left(W_j(s,t) \le y\right) = \exp\left(-\left(e^{v_j(t)} - e^{v_j(s)-(t-s)\theta}\right)e^{-y-\gamma}\right)$$

for $y \in R,$ where $\gamma = 0.5772...$, (Euler's constant) and $\{v_j(t)\}$ are deterministic functions of t such that $v_j(t) + \theta t$ is nondecreasing for all j. Moreover, $W_j(s,t)$ and $W_j(s', t')$ are independent when $(s,t) \cap (s', t') = \emptyset$. It follows readily that $v_j(t)$ has the interpretation $v_j(t) = EU_j(t)$. Tiago de Oliveira (1973) has demonstrated that when $v_j(t)$ is constant then $U_j(t)$ becomes (strictly) stationary. As demonstrated by Resnick and Roy (1990) we can express a particularly version of the autocorrelation function of the utility process (2.4) as

(2.6) $$corr\left(\exp\left(-U_j(s)\right),\exp\left(-U_j(t)\right)\right)=\exp\left(v_j(s)-v_j(t)-(t-s)\theta\right).$$

Eq. (2.6) shows that when $v_j(t)$ varies slowly over time then the autocorrelation function is close to $\exp\left(-(t-s)\theta\right)$. In other words, the parameter θ characterizes the strength of temporal persistence in the preferences. Note that from a theoretical point of view it does not matter whether we use a modified extremal process or a (standard) extremal process since the time trend (θt) cancels in utility comparisons. However, the modified extremal process formulation allows a convenient interpretation due to (2.6).

Recall that $v_j(t)+\theta t$ must be nondecreasing for all j. We shall next introduce a reparametrization that does not suffer from this restriction. This reparametrization is given by

(2.7) $$w_j(t)=v_j(t)+\log\left(\frac{\theta+v_j'(t)}{\theta}\right).$$

From (2.7) it follows that with $v_j(0)=-\infty$;

(2.8) $$\exp\left(EU_j(t)\right)=\exp\left(v_j(t)\right)=\theta\int_0^t\exp\left(w_j(\tau)-(t-\tau)\theta\right)d\tau.$$

This particular reparametrization implies that $v_j(t)+\theta t$ is increasing for any $\left\{w_j(t),t>0\right\}$. However, the main motivation behind (2.7) is that the reparametrization above is interesting for theoretical reasons. To realize this note first that when Δt is small we get from (2.5) and (2.8) that

$$P\left(W_j\left(t-\Delta t,t\right)\leq y\right)=\exp\left(-\theta\Delta t\, e^{w_j(t)-y-\gamma}+o(\Delta t)\right) \tag{2.9}$$

which shows that $w_j(t)$ has, apart form an additive term, the interpretation as the mean of "instantaneous" utility increments, $\left\{W_j\left(t-\Delta t,t\right)\right\}$. When $w_j(t)$ is independent of time (2.8) reduces to

$$\exp\left(EU_j(t)\right)=e^{w_j}\left(1-e^{-\theta t}\right). \tag{2.10}$$

Thus for large θt, constant $\left\{w_j(t)\right\}$ yield constant mean utility levels. Also from (2.8) we realize that θ is analogous to a preference time rate parameter because by (2.8), the mean utility at time t can be expressed as an integral of past weighted "instantaneous" mean utilities. Specifically, the contribution from the period s-specific systematic utility component to the current mean utility is evaluated by multiplying $\exp\left(w_j(s)\right)ds$ by the "depreciation" factor $\exp\left(-\left(t-s\right)\theta\right)$.[8] This depreciation factor accounts for the loss of memory and/or decrease in taste persistence as the time lag increases.

To clarify the interpretation further, consider the autocorrelation function (2.6) with constant $\left\{w_j(t)\right\}$. Then (2.6) reduces to

$$corr\left(\exp\left(-U_j(s)\right),\exp\left(-U_j(t)\right)\right)=\frac{1-e^{-\theta s}}{1-e^{-\theta t}}\cdot e^{-(t-s)\theta}. \tag{2.11}$$

Thus when $\min(\theta s, \theta t)$ is large, the mean utility in this case equals w_j, (apart from an additive constant) and the autocorrelation function becomes exponential.

REMARK 4: It is important to emphasize that in the discussion of the extremal process above we have made no assumptions that restricts the class of inhomogeneous extremal processes with extreme value marginals.

In the following we shall use the concept of *modified extremal process*, to mean a stochastic process which satisfies (2.4) and (2.5) with $U_j(0) = -\infty$, and with $v_j(t)$ differentiable in t for all j.

THEOREM 2: *Assume that the random utilities are independent modified extremal processes. Assume that the choice set process is constant over time. Then the choice process* $\{J(t,B), t > 0\}$ *is a Markov chain.*

The result of Theorem 2 was originally proved by Dagsvik (1983) and (1988).

Recall that by (2.4) the utility processes are Markov processes. However, utility processes with the Markov property do not usually imply that the corresponding choice process $\{J(t,B)\}$ is Markovian.[9]

3. IMPLICATIONS FOR THE STRUCTURE OF THE CHOICE PROBABILITIES

In this section we discuss the implications of the proceeding results for the structure of the choice model and its interpretation.

THEOREM 3: *Assume that the utilities are independent modified extremal processes and the choice set process is constant over time. Then for* $B \in \Im$

(3.1) $$P\left(J(t,B)=j\right)=\frac{\int_0^t e^{w_j(\tau)-(t-\tau)\theta}\,d\tau}{\sum_{k\in B}\int_0^t e^{w_k(\tau)-(t-\tau)\theta}\,d\tau},$$

and

(3.2) $$P\left(J(t,B)=j \,\middle|\, J(s,B)=i\right)=\frac{\int_s^t e^{w_j(\tau)-(t-\tau)\theta}d\tau}{\sum_{k\in B}\int_0^t e^{w_k(\tau)-(t-\tau)\theta}d\tau},$$

for $i \neq j$*, and* $P\left(J\left(t,B\right)=i \,\middle|\, J\left(s,B\right)=i\right)$ *is determined by the adding-up condition. The transition probability, given a transition out of the occupied state,* $\pi_{ij}(t)$*, equals*

(3.3) $$\pi_{ij}(t)\equiv P\left(J\left(t,B\right)=j \,\middle|\, J\left(t\text{-},B\right)\neq i\right)=\frac{e^{w_j(t)}}{\sum_{k\in B\setminus\{i\}} e^{w_k(t)}}$$

for $i \neq j$*. Moreover*

(3.4) $$Corr\left(\exp\left(-\max_{k\in B} U_k(s)\right),\exp\left(-\max_{k\in B} U_k(t)\right)\right)=\frac{\sum_{k\in B}\int_0^s e^{w_k(\tau)-(s-\tau)\theta}d\tau}{\sum_{k\in B}\int_0^t e^{w_k(\tau)-(t-\tau)\theta}d\tau}e^{-(t-s)\theta}.$$

PROOF: The results (3.1) and (3.2) follow from Dagsvik (1988) by inserting (2.8). Eq. (3.3) follows readily from (3.2). Eq. (3.4) follows from Resnick and Roy (1990).[10] *Q.E.D.*

From Theorem 3 one can easily derive the corresponding transition intensities. Dagsvik (1988) and Resnick and Roy (1990) extend the result of Theorems 2 and 3 to the case where $\{U(t), t \geq 0\}$ is a multivariate extremal process. Dagsvik considers the case where $U(t)$—at each t—has a type III multivariate extreme value distribution that is absolutely continuous. The resulting (marginal) choice probabilities at a given point in time in this case become generalized extreme value probabilities. Resnick and Roy (1990) allow $U(t)$ to have a multivariate c.d.f. that is not necessarily absolutely continuous.

The results obtained above are useful for justifying the choice of functional form of the likelihood function of observations on $\{J(\tau, B),\ \tau \leq t\}$ for a particular agent in the absence of structural state dependence. The first step in specifying an empirical model is to specify the structural parts of the model.

In empirical applications one would typically specify $w_j(t)$ as $w_j(t) = w\left(X_j(t)\right)$ where $X_j(t)$ is a vector of observable attributes specific to alternative a_j at time t and $w(\cdot)$ is a suitably chosen functional form that is known apart from an unknown vector of parameters.

Let us next consider the particular case where $\{w_j(t)\},\ j = 1, 2, ..., m,$ are constant over time i.e., $w_j(t) = w_j$. Then the transition intensity for transitions from state i to j reduces to

(3.5) $$q_{ii}(t) = -\frac{\theta\left(1 - P_i\right)}{1 - e^{-\theta t}}, \quad q_{ij}(t) = \frac{\theta e^{w_j}}{\left(1 - e^{-t\theta}\right)\sum_{k \in B} e^{w_k}} \equiv \frac{\theta P_j}{1 - e^{-\theta t}},$$

for $i \neq j$ where P_j is the probability of being in state j. Recall that the degree of serial correlation in the indirect utility can be measured by θ. Specifically, when θ is large there is weak serial correlation (provided θt is large) while when θ is close to zero tastes are strongly correlated over time. Moreover, (3.5) shows that the transition intensities are stationary when t is large. However, when t is small then the transition intensities given by (3.5) depend on time. This is due to the fact that in the beginning of a choice process the length of the choice history (age) will influence the strength of the serial correlation of the utilities.

Let us finally consider the probability distribution of the holding time in state i. Let $T_i(s)$ be the holding time in state i given that a_i was chosen at time s. Since the choice process is an inhomogeneous Markov chain we have that

(3.6) $$P\left(T_i(s) > y\right) = \exp\left(\int_s^{s+y} q_{ii}(x)\,dx\right).$$

From (3.5) we realize that in the stationary case the hazard function, $\theta\left(1-P_i\right)$, belongs to the *proportional hazard family*.[11] However, the hazard function $\theta\left(1-P_i\right)$ does not have a simple Cox type of structure as a loglinear function of attributes specific to alternative i, but depends on the attributes of all the alternatives in the choice set.

From (3.2) we realize that when $i \neq j$ then the transition probability of going from i to j does *not* depend on i. This, however, does not necessarily mean that the corresponding aggregate transition probabilities are independent of i. To realize this consider again the case where $w_j(t)$ is constant over time, but depend on individual characteristics x (say). Then (3.2) reduces to

$$\frac{\left(1-e^{-(t-s)\theta}\right)P_j(x)}{1-e^{-t\theta}}$$

when $i \neq j$, where $P_j(x)$ is the probability of being in state j conditional on x. However, the corresponding *aggregate* transition probability equals

$$\frac{\left(1-e^{-(t-s)\theta}\right)E\left(P_i(x)P_j(x)\right)}{\left(1-e^{-t\theta}\right)E\,P_i(x)}$$

where expectation is taken with respect to x. Evidently, this expression is *not* necessarily independent of i.

4. EXTENDING THE MODEL TO ALLOW FOR STATE DEPENDENCE

So far we have only discussed the functional form of the choice probabilities of $\left\{J\left(t,B\right)\right\}$ when there is no structural dependence from past choice experience. The question now arises how the particular functional form that follows should be modified in the presence of state dependence. The interpretation of this setting is that the agent has myopic preferences and he does not take into account how past and current experience affect future preferences. Notice first that when the utility processes are altered by the choice history the structural terms of the utility processes become endogeneous. This is so because the structural terms of the utility processes become dependent on past choices, and consequently they will depend on past realizations of the utility processes.

Let $h(t)$ denote the choice history prior to t. For expository simplicity, consider the discrete time case. One natural way of introducing state dependence is to assume that the utility processes are independent modified experience-dependent extremal processes defined by

(4.1) $$U_j(t) = \max\left(U_j(t-1) - \theta, W_j\left(t, h(t)\right)\right)$$

where $W_j\left(t, h(t)\right)$ is a random variable with distribution

(4.2) $$P\left(W_j\left(t, h(t)\right) \le u \,\middle|\, U(t-1)\right) = \exp\left(-\exp\left(w_j\left(t, h(t)\right) - u\right)\right)$$

and where $w_j\left(t, h(t)\right)$ is a parametric function of the attributes of alternative j and past choice experience.

In Dagsvik (2000) it is demonstrated that under (4.2) the transition probabilities are modified versions of the ones given in Theorem 3 where the modification consists in replacing $\left\{w_j(t)\right\}$ by $\left\{w_j\left(t, h(t)\right)\right\}$ and replacing the integrals by sums. In other words, we can treat the "choice history variable", $h(t)$, as if it were exogenous. Now provided the choice set B contains at least 3 alternatives, including a_1, a_2 and a_j, (3.3) (with $w_j(t)$ replaced by $w_j\left(t, h(t)\right)$) implies that $w_j\left(t, h(t)\right) - w_1\left(t, h(t)\right)$ is nonparametrically identified since it is determined by $\pi_{2j}\left(t, h(t)\right) / \pi_{21}\left(t, h(t)\right)$. Thus, if one believes that IIIA represents a reasonable behavioral assumption (under the absence of state dependence) the modeling framework developed in this paper allows one to identify state dependence effects. We refer to Heckman (1978), (1981a), (1981b), (1991) and Keane (1997) for a further discussion of this issue.

To express the corresponding result in continuous time it is necessary to apply the max spectral representation of the extremal process, cf. Resnick and Roy (1990). This is, however, beyond the scope of this paper.

5. CONCLUSIONS

In this paper we have considered the problem of extending the IIA Axiom to the intertemporal setting. It is demonstrated that a particular extension of Luce IIA axiom implies a random utility model where the utilities are extremal processes. In myopic settings with no state dependence effects and choice sets that are nondecreasing over time this model has the Markov property with a particular structure of the transition probabilities. Finally, we discuss how the choice model can be extended to allow for structural state dependence.

Although the modeling framework discussed there is based on a static setting it can nevertheless be applied to analyze particular intertemporal discrete/continuous choice problems under perfect foresight and two stage budgeting. An example of this approach is provided by Heckman and MaCurdy (1980) where the choice of working versus not working is reduced to a static one conditional on a fixed effect representation of the initial marginal utility of wealth. (See also Blundell and MaCurdy (1999) for a review of this type of approach.) Although Heckman and MaCurdy only consider binary discrete choice this technique could readily be extended to cover cases with multinomial choices (corner solutions), see Dagsvik (2000).

Microeconometric Research Division, Statistics Norway, P.B. 8131 Dep., 0033 Oslo, Norway; E-mail: john.dagsvik@ssb.no

APPENDIX

Below we shall draw on the properties of the multivariate extreme value distribution. To this end we start by listing some of the properties of the bivariate extreme value distribution. Let $F(x,y)$ be a standardized type III bivariate extreme value distribution. Then for any $z \in R$,

(A.1) $$\log F(x,y) = e^{-z} \log F(x-z, y-z),$$

(cf. Resnick, 1987). In general, the multivariate extreme value distribution is not absolutely continuous with respect to the Lebesgue measure. However, we have the following result:

LEMMA 1: *Let $F(x,y)$ be a bivariate (type III) extreme value distribution. Then $-\log F(-x,-y)$ is convex and the left and right derivatives, $\partial F_{\pm}(x,y)/\partial x$ and $\partial F_{\pm}(x,y)/\partial y$, exist and are non-decreasing.*

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

$$\Delta = \{z \in R_+^2 : z_1^2 + z_2^2 = 1\}$$

such that

$$L(x,y) = \int \max(z_1 e^x, z_2 e^y)\, \mu(dz_1, dz_2).$$

Since z_1e^x and z_2e^x are convex functions in x it follows that $L(x,y)$ is convex. Since $L(x,y)$ is convex the left and right derivatives of $F(x,y)$ exist. (See for example Kawata, Theorem 1.11.1 p. 27.) *Q.E.D.*

From now on the notion "derivative", will mean the (first order) right derivative.

Let $F(x,y)$ be a bivariate (type III) extreme value distribution and let

$$\varphi(x) = -\log F(-x,0).$$

Then it follows immediately from (A.1) that

(A.2) $$\log F(x,y) = -e^{-y}\varphi(y-x)$$

which implies that

(A.3) $$\partial_2 \log F(-x,0) = \varphi(x) - \varphi'(x)$$

where ∂k means the derivative with respect to component *k*. The relations above will be useful in the proof below.

PROOF OF THEOREM 1: We assume $n=2$ since the general case is completely analogous. Let $B(s)=B_1=\{i,j\}$, $s<t$, and $B(t)=\{i,j,k\}$. Let $F_b(x,y)$ be the c.d.f. of

$\left(U_b(s),U_b(t)\right)$, $s<t$. By Axiom A5 $F_b(x,y)$ is a bivariate extreme value (type III) distribution. Note that since $\{U_b(t), t\geq 0\}$ is assumed continuous in probability it follows that $F_b(x,y)$ is continuous in (x,y). Let

$$G_b(x,y) = -\log F_b(x,y).$$

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

(A.4)
$$\begin{aligned}
&P\left(U_i(s)>U_j(s),U_j(t-)>U_i(t-),U_k(t)>\max\left(U_i(t),U_j(t)\right)\right)\\
&=P\left(U_i(s)>U_j(s),U_k(t)>U_j(t)>U_i(t)\right)\\
&=\int_{\substack{y_1<y_2<y_3\\ x_1>x_2}} F_i\left(dx_1,dy_1\right)F_j\left(dx_2,dy_2\right)F_k\left(\infty,dy_3\right)\\
&=\int\left(1-F_k\left(\infty,y\right)\right)F_j\left(x,dy\right)F_i\left(dx,y\right).
\end{aligned}$$

Since the marginal distribution of $U_b(t)$ is type III extreme value (Proposition 2) we can write $F_b(\infty,y)$ as

(A.5)
$$F_b(\infty,y) = \exp\left(-m_b e^{-y}\right)$$

where $\log m_b = EU_b(t) - 0.5772$. By Lemma 1 the first order left and right derivatives of $G_b(x,y)$ exist. Since F by assumption is a bivariate extreme value distribution it follows from (A.1) that for $z\in R$

(A.6) $$F_j\left(x,dy\right)F_i\left(dx,y\right)=\left[\exp\left(-e^{-y}\left(G_i(x-y,0)\right)+G_j(x-y,0)\right)\right]$$
$$\cdot\, e^{-2y}\,\partial_1 G_i(x-y,0)\,\partial_2 G_j(x-y,0)\,dx\,dy.$$

Let $\varphi_b(x)=G_b(-x,0)$. From (A.2) and (A.3) it follows that

$$G_b(x,y)=e^{-y}\,G_b(x-y,0)=e^{-y}\,\varphi_b(y-x)$$

and

(A.7) $$\partial_2 G_b(-x,0)=-\varphi_b(x)+\varphi_b'(x).$$

By Lemma 1, $\varphi_b(x)$ is convex and therefore has derivatives that are non-decreasing. From (A.4), (A.5), (A.6) and (A.7) it follow after the change of variable $x=-u+y$, that

(A.8) $$P\left(U_i(s)>U_j(s),U_k(t)>U_j(t)>U_i(t)\right)$$
$$=\int_{R^2}\left(1-\exp\left(m_k e^{-y}\right)\right)\left[\exp\left(-e^{-y}\left(\varphi_i(u)+\varphi_j(u)\right)\right)\right]\varphi_i'(u)\left(\varphi_j(u)-\varphi_j'(u)\right)e^{-2y}du\,dy.$$

Due to the fact that for any $c>0$

$$\int_R \exp\left(-c\,e^{-y}\right)e^{-2y}dy=\frac{1}{c^2},$$

(A.8) reduces to

(A.9)
$$P\left(U_i(s) > U_j(s), U_k(t) > U_j(t) > U_i(t)\right)$$
$$= \int_R \frac{\varphi_i'(u)\left(\varphi_j(u) - \varphi_j'(u)\right) du}{\varphi_{ij}(u)^2} - \int_R \frac{\varphi_i'(u)\left(\varphi_j(u) - \varphi_j'(u)\right) du}{\left(\varphi_{ij}(u) + m_k\right)^2},$$

where $\varphi_{ij}(u) = \varphi_i(u) + \varphi_j(u)$. Now Axiom A1 and (A.4) imply that

(A.10)
$$P\left(U_i(s) > U_j(s), U_k(t) > U_j(t) > U_i(t)\right)$$
$$= P\left(U_i(s) > U_j(s), U_j(t-) > U_i(t-), U_k(t) > \max\left(U_j(t), U_j(t)\right)\right)$$
$$= P\left(U_k(t) > \max\left(U_i(t), U_j(t)\right)\right) P\left(U_i(s) > U_j(s), U_j(t) > U_i(t)\right)$$
$$= \frac{m_k}{m_i + m_j + m_k} \cdot P\left(U_i(s) > U_j(s), U_i(t) < U_j(t)\right).$$

From (A.9) we obtain, by letting $m_k \to \infty$, that

(A.11)
$$P\left(U_i(s) > U_j(s), U_i(t) < U_j(t)\right) = \int_R \frac{\varphi_i'(u)\left(\varphi_j(u) - \varphi_j'(u)\right) du}{\varphi_{ij}(u)^2}.$$

Hence, (A.9), (A.10) and (A.11) imply that

(A.12)
$$\frac{m_i + m_j}{m_i + m_j + m_k} \int_R \frac{\varphi_i'(u)\left(\varphi_j(u) - \varphi_j'(u)\right) du}{\varphi_{ij}(u)^2} = \int_R \frac{\varphi_i'(u)\left(\varphi_j(u) - \varphi_j'(u)\right) du}{\left(\varphi_{ij}(u) + m_k\right)^2}.$$

Suppose now that $x = r \geq -\infty$, is the largest point at which $\varphi_i'(x) + \varphi_j'(x) = 0$. Then since $\varphi_b'(x)$ is nondecreasing and non-negative it must be true that $\varphi_i'(x) = \varphi_j'(x) = 0$ for $x \leq r$. As a consequence the mapping $\psi_{ij} : R_+ \to [r, \infty)$, defined by

(A.13) $$z = \varphi_{ij}\left(\psi_{ij}(z)\right) - \varphi_{ij}(r)$$

exists, is invertible and has (right) derivatives everywhere on R_+. By change of variable

$$u \to \psi_{ij}^{-1}(u) = z$$

(A.12) takes the form

(A.14) $$\frac{m_i + m_j}{m_i + m_j + m_k} \int_0^\infty \frac{f_{ij}(z)\,dz}{\left(\varphi_{ij}(r) + z\right)^2} = \int_0^\infty \frac{f_{ij}(z)\,dz}{\left(\varphi_{ij}(r) + z + m_k\right)^2},$$

where

(A.15) $$f_{ij}(z) = \varphi_i'\left(\psi_{ij}(z)\right)\psi_{ij}'(z)\left(\varphi_j\left(\psi_{ij}(z)\right) - \varphi_j'\left(\psi_{ij}(z)\right)\right).$$

But the right hand side of (A.14) is a *generalized Stieltjes transform* of $f_{ij}(\cdot)$ (see Widder, 1938), evaluated at $\varphi_{ij}(r) + m_k$. The generalized Stieltjes transform is well defined provided $f_{ij}(z)$ is integrable and the integral (A.14) exists. The generalized Stieltjes transform of a function determines the function uniquely almost everywhere. Thus, due to the uniqueness property of the generalized Stieltjes transform, (A.14) implies that $f_{ij}(\cdot)$ must be constant almost everywhere for $z \geq 0,$ since the left hand side of (A.14) is the generalized Stieltjes transform of a constant. As a consequence, we must have that $m_i + m_j = \varphi_{ij}(r).$ From the definition of $\psi_{ij}(z)$ we get

(A.16) $$1 = \varphi'_{ij}\left(\psi_{ij}(z)\right)\psi'_{ij}(z).$$

Hence, (A.15) and (A.16) with $u = \psi_{ij}(z),$ yield

(A.17) $$\varphi'_i(u)\left(\varphi_j(u) - \varphi'_j(u)\right) = \varphi'_{ij}(u)C_{ij}$$

for $u > r,$ where C_{ij} is a constant. Similarly we get, by interchanging i and j in the demonstration above, that

(A.18) $$\varphi'_j(u)\left(\varphi_i(u) - \varphi'_i(u)\right) = \varphi'_{ij}(u)C_{ji}$$

for $u > r,$ where C_{ij} is another constant. By subtracting (A.18) from (A.17) we get

$$\varphi'_{ij}(u)\varphi_j(u) - \varphi'_j(u)\varphi_{ij}(u) = \varphi'_{ij}(u)\left(C_{ij} - C_{ji}\right)$$

which, when dividing by $\varphi_{ij}(u)^2$ becomes equal to

(A.19) $$\frac{\varphi_j'(u)\,\varphi_{ij}(u)-\varphi_j(u)\,\varphi_{ij}'(u)}{\varphi_{ij}(u)^2}=\frac{\varphi_{ij}'(u)\left(C_{ij}-C_{ji}\right)}{\varphi_{ij}(u)^2}.$$

Next, integrating both sides of (A.19) yields

$$\frac{\varphi_j(u)}{\varphi_{ij}(u)}=\frac{C_{ij}-C_{ji}}{\varphi_{ij}(u)}+d_1$$

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

(A.20) $$\varphi_j(u)=C_{ij}-C_{ji}+\varphi_{ij}(u)\,d_1$$

for $u>r$. By inserting (A.20) into (A.18) we get

$$\varphi_{ij}'(u)\left(\varphi_i(u)-\varphi_i'(u)\right)d_1=\varphi_{ij}'(u)\,C_{ji}$$

for $u>r$, which, since $\varphi_{ij}'(u)>0$, is equivalent to

(A.21a) $$\varphi_i(u)-\varphi_i'(u)=C_{ji}/d_1.$$

Similarly, it follows that

(A.21b) $$\varphi_j(u)-\varphi_j'(u)=C_{ij}/d_2$$

for $u > r$. Eq. (A.21a,b) are first order differential equations which have a solution of the form

(A.22) $$\varphi_b(u) = \alpha_b + \beta_b e^u$$

for $u > r \geq -\infty, b = i, j$. Since $\varphi_b'(u) = 0$ for $u \leq r$, and $\varphi_b(u)$ is continuous we get from (A.22) that

(A.23) $$\varphi_b(u) = \alpha_b + \beta_b e^r$$

for $u \leq r$. As a consequence it must be true that (almost everywhere)

(A.24) $$G_b(x, y) = e^{-y}\varphi_b(y - x) = \alpha_b e^{-y} + \beta_b \exp\left(-\min(x, y - r)\right).$$

From (A.24) we obtain that for $s < t$

(A.25) $$P\left(U_b(t) \leq y \,\middle|\, U_b(s) = x\right) = 0$$

when $y < x + r$, and

(A.26) $$P\left(U_b(t) \leq y \,\middle|\, U_b(s) = x\right) = P\left(U_b(t) \leq y\right)$$

when $y \geq x + r$. Eq. (A.25) means that $\{U_b(t)\}$ is non-decreasing with probability one. Eq. (A.26) means that conditional on $U_b(t) > U_b(s)$, $U_b(t)$ is stochastically independent of $U_b(s)$. But then we must have that $\{U_b(t)\}$ is equivalent to the utility process defined by

(A.27) $$U_b(t) = \max\left(U_b(s), W_b(s,t)\right) + r$$

where $W_b(s,t)$ is extreme value distributed and independent of $U_b(s)$. Since $U_1(t) - U_2(t)$ is independent of r for any t we may without loss of generality choose $r = 0$. But since (s,t), with $s < t$, were arbitrarily chosen points in time, (A.27) defines the (inhomogeneous) extremal process (cf. Dagsvik, 1988) which was to be proved. *Q.E.D.*

REFERENCES

AN, M.Y. (1995): "Econometric Analysis of Sequential Discrete Choice Models," Manuscript, Department of Economics, Duke University.

BLUNDELL, R. AND T. MACURDY (1999): "Labor Supply: A Review of Alternative Approaches," in *Handbook of Labor Economics*, ed. by O. Ashenfelter and D. Card. Elsevier Science B.V.

DAGSVIK, J.K. (1983): "Discrete Dynamic Choice: An Extension of the Choice Models of Luce and Thurstone," *Journal of Mathematical Psychology*, 27, 1-43.

— (1988): "Markov Chains Generated by Maximizing Components of Multi-dimensional Extremal Processes," *Stochastic Processes and Their Applications*, 28, 31-45.

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

— (2000): "A Behavioral Characterization of Intertemporal Random Utility Models," Mimeo, Statistics Norway.

DE HAAN, L. (1984): "A Spectral Representation for Max-stable Processes," *Annals of Probability*, 12, 1194-1204.

HAUSMAN, J. AND D. MCFADDEN (1984): "A Specification Test for the Multinomial Logit Model," *Econometrica*, 52, 1219-1240.

HECKMAN, J.J. (1978): "Simple Statistical Models for Discrete Panel Data Developed and Applied to Test the Hypothesis of True State Dependence against the Hypothesis of Spurious State Dependence," *Annales de l'INSEE*, 30-31, 227-269.

— (1981a): "Statistical Models for the Analysis of Discrete Panel Data,". in *Structural Analysis of Discrete Data*, ed. by C.F. Manski and D. McFadden. Cambridge: MIT Press.

— (1981b): “Heterogeneity and State Dependence,” in (eds.), *Studies in Labor Markets*, ed. by S. Rosen. Chicago: Univ. of Chicago Press.

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

HECKMAN, J.J. AND T. MACURDY (1980): “A Life Cycle Model of Female Labour Supply,” *Review of Economic Studies*, 46, 47-74.

KAWATA, T. (1972): *Fourier Analysis in Probability Theory*. New York: Academic Press.

KEANE, M.P. (1997): “Modeling Heterogeneity and State Dependence in consumer Choice Behavior,” *Journal of Business & Economic Statistics*, 15, 310-327.

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

LUCE, R.D. AND P. SUPPES (1965): “Preference, Utility, and Subjective Probability,” in *Handbook of Mathematical Psychology*, III, 249-410, ed. by R.D. Luce, R.R. Bush, and E. Galanter. New York: Wiley.

OLSEN, R.J., D. ALTON SMITH, AND G. FARKAS (1986): “Structural and Reduced-Form Models of Choice among Alternatives in Continuous Time: Youth Employment under a Guaranteed Jobs Program,” *Econometrica*, 54, 375-394.

RESNICK, S. (1987): *Extreme Value, Regular Variation and Point Processes*. New York: Springer Verlag.

RESNICK, S. AND R. ROY (1990): “Multivariate Extremal Processes, Leader Processes and Dynamic Choice Models,” *Advances in Applied Probability*, 22, 309-331.

STRAUSS, D. (1979): “Some Results on Random Utility Models,” *Journal of Mathematical Psychology*, 17, 32-35.

TIAGO DE OLIVEIRA, J. (1973): “An Extreme Markovian Stationary Process,” *Proceedings of the Fourth Conference in Probability Theory*, Brasov: Acad. Romania, 217-225.

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

WIDDER, D.V. (1938): "The Stieltjes Transform," *Transactions of the American Mathematical Society*, 43, 7-60.

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

Footnotes:

[1] Thanks to Rolf Aaberge, Leif Andreassen, Tor Jakob Klette, Tom Kornstad, Pedro de Lima, Chris Flinn, John Rust and Steinar Strøm for criticism and comments and Anne Skoglund for technical assistance and correction of errors. I also thank participants in workshops at the University of California, Berkeley, University of Chicago and University of Wisconsin for comments on earlier versions of the paper. I am particularly grateful for all help from Richard Blundell through the revision and editing process.

[2] Although IIA is a strong assumption, and clearly does not hold in situations in which alternatives are perceived as dissimilar, it should be remembered that due to unobserved population heterogeneity it may very well be the case that it holds conditional on relevant individual characteristics but fail to hold when individual characteristics are not properly controlled for.

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

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

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

[4] In Dagsvik (1983), p. 11, a particular behavioral assumption (Axiom 1) was proposed. This axiom is weaker than A1. Axiom 1 makes a statement analogous to IIA for choice careers $J(t_1) = j_1, J(t_2) = j_2, J(t_3) = j_3$ at time epochs $t_1 < t_2 < t_3$, where j_1, j_2 and j_3 are all different.

[5] Recall that the notation $B \backslash A$ means the set of elements which belong to B but not to $A \cap B$.

[6] In a Bernoulli process there is no serially dependence.

[7] Recall that the type III extreme value distribution has the form $\exp\left(-ae^{-bx}\right)$ where $a > 0, b > 0$ are constants (cf. Resnick, 1987).

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

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

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

$$EU_j^* = e^{w_j}.$$

[9] It is not known whether the class of extremal processes is the only one that yields the Markov property of the choice model in continuous time.

[10] Similarly to Dagsvik (1983) it can also be proved that the autocorrelation function in (2.16) has the interpretation as

$$corr\left\{\max_{k\in B(s)} U_k(s), \max_{k\in B(t)} U_k(t)\right\} = \rho\left(\frac{\sum_{k\in B(s)} \int_0^s e^{w_k(\tau)-(s-\tau)\theta} d\tau}{\sum_{k\in B(t)} \int_0^t e^{w_k(\tau)-(t-\tau)\theta} d\tau} \cdot e^{-(t-s)\theta}\right)$$

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

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

(cf. Tiago de Oliveira, 1973).

[11] See An (1995) for a justification of the proportional hazard family derived within a dynamic choice setting under uncertainty.

List of symbols

S

m

a_j

$U_j(t), U_j(s)$

t

$v_j(t)$

$\varepsilon_j(t)$

$J(t, B(t))$

$J(t, B)$

$B(t)$, B, $B(s)$

$\Im$

P

$\forall$

$\backslash$

$\varnothing$

Z, $Z_j(t)$

$\tilde{v}(\cdot)$, $\tilde{v}(Z_j(t))$

$v_j(t)$

K

x (ex)

$\in$ (element of)

$\notin$ (not element of)

$\geq$

θ

$-\infty, \infty$

∞

R

γ

$W_j(s,t), W_j(s',t')$

$\cap$

$\int$

$w_j(\tau)$

τ

Δ

$\pi_{ij}, \pi_{ij}(t)$

$\tau-$

$w(X_j(t))$

$X_j(t)$

P_i, P_j, $P_i(x)$

$q_{ii}(t)$

$T_i(s)$

$h(t)$

y, z, u

L

F, F_j

G, G_j

$\partial, \partial_j, \partial_1, \partial_2$

$\varphi_b(x), \varphi_b, \varphi_j'$

m_j

ψ_{ij}, ψ_{ij}'

$\varphi_{ij}, \varphi_{ij}'$

ψ_{ij}^{-1}

f_{ij}

α_b

β_b

r

$\|\cdot\|$ \`

η_1, η_2

$\exists$

$Y(t)$

$J(t_1)$

κ

Γ

ρ

π