Dynamic models of choice behaviour : some fundamentals and trends

HARRY TIMMERMANS AND ALOYS BORGERS

DYNAMIC MODELS OF CHOICE BEHAVIOUR: SOME FUNDAMENTALS AND TRENDS

1. INTRODUCTION

In this chapter we will discuss some models of dynamic decision-making and choice behaviour. The term dynamic indicates that we are interested in choice behaviour over time, i.e. in possible changes in choices. The discussion will be restricted to three general types of modelling approaches. First, stochastic models of buying behaviour will be discussed. This is followed by a summary of newly developed variety-seeking models. Finally a brief summary is given of some of the latest developments in the field of dynamic discrete choice models. It should be noted that these models constitute only a small part of this rapidly growing field of research and that the structure of this section is rather arbitrary. Other reviews can be found in Halperin and Gale (1984), Halperin (1985) and Hensher and Wrigley (1984).

2. STOCHASTIC MODELS OF BUYING BEHAVIOUR

The class of stochastic models of buying behaviour consists of brand choice models and purchase incidence models. Brand choice models predict which choice alternative will be chosen given that a choice is made at a particular point in time. Purchase incidence models predict when an alternative will be chosen or how many alternatives will be chosen in a specified interval of time.

Several models belonging to the class of stochastic buying behaviour will be discussed in the following subsections. For a more detailed survey, we refer to Massy, Montgomery and Morrison (1970).

2.1. *Bernoulli Models*

Perhaps the most simple model of dynamic choice behaviour is the Bernoulli model. It is based on the assumption that the probability of choosing alternative i is constant over time. It implies that the past history of the process has no effect on the choice probabilities. Hence, the model may be expressed as:

$$p_t = p \tag{1}$$

where p_t is the probability that a particular alternative is chosen at time t;
p is the initial probability of choosing the alternative.

Another limiting property of the Bernoulli model is that it assumes homogeneity. That is, it is assumed that the choice probabilities apply to all individuals. Nevertheless the model has been frequently

J. Hauer et al. (eds.), Urban Dynamics and Spatial Choice Behaviour, 3–26.

used with success, especially in a marketing context. Burnett (1975) applied the Bernoulli model in a study of spatial shopping behaviour using panel data. The model did not perform very well though.

2.2. *The Compound Beta Bernoulli Model*

Perhaps the most important limitation of the standard Bernoulli model is its homogeneity assumption. The compound beta Bernoulli model has been developed to relax this assumption. The model still assumes that every individual in the population has a constant probability p of choosing an alternative and a probability of $(1-p)$ of choosing another alternative, but the homogeneity assumption is replaced by assuming that p has a beta distribution over individuals in the population. This beta distribution has the form:

$$b(p) = \begin{cases} \dfrac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\,\Gamma(\beta)}\, p^{\alpha-1}(1-p)^{\beta-1}, & 0 < p < 1 \\ 0, \text{ otherwise} & \end{cases} \tag{2}$$

where $\Gamma(\cdot)$ is the gamma function;
α and β are parameters to be estimated $(\alpha,\ \beta > 0)$.

2.3. *Markov Models*

Markov models have been used to study the dynamics of choice behaviour. Especially the first-order models have been applied frequently. These models are typically based on deriving a transition probability matrix which expresses the probability that alternative j will be chosen at time t+1 given that alternative i has been chosen at time t. Consequently, these transition or switching probabilities are independent from the choices at times t-1, t-2, ... Moreover, in the conventional models homogeneity and stationarity is assumed: the transition probabilities apply to all individuals in the population and the transition probability matrix is independent of t. Given these assumptions, choice behaviour and market shares at some future point at time t^* can be calculated easily by raising the transition probability matrix to the power t^* and multiplying this matrix by an initial state vector. Likewise, the vector of steady state probabilities can be calculated in a straightforward manner.

This standard Markov model has been very popular in marketing science. A geographical example is provided in Burnett (1974, 1978) and Crouchley, Davies and Pickles (1982), although it should be noted that Markov chain analysis as such is a well-known technique in geography which has been used for a variety of purposes.

It is evident that the first-order Markov chain is based on some rigorous assumptions: first order, homogeneity and stationarity. Hence, several authors have attempted to develop more sophisticated Markov models which relax such assumptions. Higher-order Markov models have been developed to incorporate the effect of choice behaviour at times t-1, t-2, etc. on choice behaviour at time t+1. Non-stationary models have been proposed to make the transition probability matrix dependent upon the time period. Heterogeneity has been introduced by segmentation of the population into groups. Some of these developments are well known in geographical research. In particular, such sophisticated Markov models have been used in studying migration (e.g. Ginsberg, 1971, 1972, 1973, 1978, 1979).

2.4. *Brand Loyal Models*

Brand loyal models have been developed by marketing scientists, but in a geographic context they have been called place loyal models. Basically, this model is a compound Markov model. It is based on the following assumptions: each individual's choice behaviour can be described in terms of a first-order process; the parameter describing this process has a probability distribution $b(p)$ among the individuals in the population. In particular, it is assumed that the first-order 0-1 process has the transition matrix:

	1	0
1	p	$1-p$
0	kp	$1-kp$

(3)

where k is a constant for each individual ($0 < k < 1$).

$b(p)$ is beta distributed, although any arbitrary probability density is allowed.

2.5. *Last Purchase Loyal Models*

An alternative to the brand loyal model is the last purchase loyal model which has the following transition matrix:

	1	0
1	p	$1-p$
0	$1-kp$	kp

(4)

Apart from the structure of the transition matrix, this model is based on the same assumptions as the brand loyal model.

2.6. *Linear Learning Models*

Linear learning models were originally developed in psychology for the purpose of describing data from laboratory experiments on adaptive behaviour. However, they have also played an important role in the study of brand choice behaviour. In a geographical context, the model has hardly been used, an exception being the study by Burnett (1977) on spatial shopping behaviour. Actually, she found that the linear learning model outperformed both the Bernoulli model and the Markov model.

The basic assumption of the linear learning model is that choice probabilities at time t+1 are a linear function of the choice probabilities at time t. Further, the model assumes quasi-stationarity in the sense that the parameters of the model do not change over short periods of time and that all individuals exhibit adaptive behaviour that can be described by a single set of parameters. The model may be expressed as follows:

$$p_{t+1} = \begin{cases} \alpha + \beta + \lambda p_t & \text{if the choice alternative is chosen at time } t \\ \alpha + \lambda p_t & \text{if the choice alternative is not chosen at time } t \end{cases} \qquad (5)$$

where α, β are parameters denoting the intercept of the rejection operator and the difference in intercept between the acceptance and the rejection operator;
λ is the slope of both operators.

The expectation of p_t for $t > 0$ is equal to:

$$E(p_t | p_0) = \left[\alpha + \alpha(\beta + \lambda) + \alpha(\beta + \lambda)^2 + \ldots + \alpha(\beta + \lambda)^{t-1}\right] + (\beta + \lambda)^t \, p_0 \qquad (6)$$

This equation can be used to predict dynamic choice behaviour. The linear learning model can be considered as a generalization of both the zero order Bernoulli and the first-order Markov models. If $\lambda = 0$, the linear learning model collapses into a Markov model; if $\alpha = \beta = 0$ and $\lambda = 1$, we have a Bernoulli model.

2.7. *Purchase Incidence Models*

Again, these models were originally developed by marketing scientists in the fifties. The main differences between previously discussed dynamic models and purchase incidence models is that the latter are primarily concerned with the problem of *when* an alternative will be chosen or, equivalently, how many choices will be made in a specified interval of time, whereas the former are concerned with predicting which alternative will be chosen given that a choice is made at a particular point in time.

A well-known model in this class is Ehrenberg's negative binomial model. It is based on the following assumptions: the average number of purchases of a particular alternative is constant over time periods; the purchases of an individual over successive equal-length time periods can be described by the Poisson distribution:

$$\Pr\{N_t = k|\mu\} = f(N_t|\mu) = \exp(-\mu)\mu^k / k!, \qquad k \geq 0 \qquad (7)$$

where N_t is the total number of purchases during a fixed interval of time;
μ is the average purchase rate,

and, finally, that these average purchase rates are distributed over the individuals according to a gamma distribution with parameters α and β:

$$f(\mu) = \beta \exp(-\beta\mu)(\beta\mu)^{\alpha-1} / \Gamma(\alpha), \qquad \mu > 0 \tag{8}$$

Given these assumptions, the aggregate distribution of purchase events follows the following negative binomial model:

$$\begin{aligned} f(N_t) &= \int_0^\infty f(N_t|\mu)f(\mu)d\mu \\ &= \left[\frac{\beta}{1+\beta}\right]^\alpha \left[\frac{(N_t+\alpha)}{N_t\;\Gamma(\alpha)}\right]\left[\frac{1}{1+\beta}\right]^{N_t} \end{aligned} \tag{9}$$

Other interesting measures, such as the incidence of repeat buying, average purchase frequency and market penetration can be derived in a straightforward manner (see Ehrenberg, 1972 for details).

The basic model has been extended in a number of important ways to relax its assumptions. Chatfield, Ehrenberg and Goodhardt (1966) have shown that the distribution of the total number of purchases in a fixed time interval for individuals who purchase the item at least once during the interval can be approximated by the logarithmic series distribution:

$$f(N_t|N_t>0) = -q^{N_t} / N_t \ln(1-q) \tag{10}$$

where q is the parameter of the distribution.

Other distributions, such as the zero-truncated negative binomial (Zufryden, 1977) and the geometric model have been used.

The original model is based on the assumption that the number of purchases of a particular brand by a single individual in equal-length successive time periods are independent and follow a Poisson distribution, implying that inter-purchase times are exponentially distributed. Some authors have argued that such inter-purchase times are more regular and they have introduced the Erlang 2 distribution to describe inter-purchase times (Herniter, 1971; Jeuland *et al.*, 1980). Empirical evidence however has not substantiated this claim: only a small improvement in fit has been obtained by using the resulting negative binomial distribution. In a geographical context, Dunn, Reader and Wrigley (1983) also found that the negative binomial model holds for the majority of their respondents.

The negative binomial model has been applied successfully in marketing science (see Ehrenberg, 1968, 1972) for brand choice behaviour. Wrigley (1980) was the first geographer to use this model in a study of purchasing patterns at particular store types. In a follow up study, the model was successfully applied to purchasing at individual stores in Cardiff (see Wrigley and Dunn, 1984a; Dunn, Reader and Wrigley, 1983).

Another important development is the generalization of the model to more brands or stores. This extension is the Dirichlet model which specifies probabilistically how many purchases each consumer makes in a time-period and which brand/store is chosen on each occasion (Goodhardt, Ehrenberg and Chatfield, 1984), The model combines aspects of purchase incidence and choice aspects and is based on the following assumptions:

- individuals' choice probabilities are constant over time and independent over successive purchases, implying that the number of purchases of each alternative an individual makes in a sequence of purchases can be modelled by a multinomial model;
- these choice probabilities vary across individuals according to a Dirichlet distribution;
- successive purchases of an individual are independent with a constant mean rate, implying that the number of purchases made in each of a succession of equal-length time periods follows a Poisson distribution;
- the mean purchasing rates vary among individuals according to a gamma distribution;
- the choice probabilities and average purchase frequencies of different individuals are distributed independently over the population.

The model itself is then obtained by mixing these multinomial, Dirichlet, Poisson and gamma distributions. The model has worked remarkably well in marketing science (Goodhardt *et al*., 1984) but also for predicting brand purchases within store groups (Kau and Ehrenberg, 1984) and multistore purchase patterns within individual stores (Wrigley and Dunn, 1984b, 1984c).

A similar model was developed by Jeuland, Bass and Wright (1980) but they replaced the Poisson distribution by the Erlang 2 distribution, which results in a multiple hyper-geometric model. This model worked well in a study of purchases of cooking oil. Yet another model has been advanced by Zufryden (1977). His model is also based on the Erlang and gamma distribution of purchase incidence and the assumption of independence between choice behaviour and purchase incidence behaviour, but he uses a linear model with purchase probabilities varying over the population according to a beta distribution rather than a multinomial distribution. His model also performed extremely well (Zufryden, 1978).

All of the above models do not incorporate explanatory variables. Consequently, the effects of managerial or planning decisions on behaviour cannot be assessed. In order to circumvent this disadvantage a number of authors have proposed models which basically add explanatory variables to the models discussed in this section. For example, Jones and Zufryden (1980) proposed a logit model to explain brand choice probability as a function of purchase explanatory variables. Following Ehrenberg, they assume the negative binomial distribution to describe the product class purchase distribution over the population. The probability of choosing a particular alternative given that the class purchase is being made is then modelled by a logit model assuming heterogeneity among individuals with respect to these choice probabilities (beta distribution), independence from past purchase outcomes and time invariance. Examples of applications of this approach can be found in Jones and Zufryden (1980, 1982). In a similar vein, Paull (1978) used a polytomous logit regression approach to predict discrete purchase quantities in a generalized negative binomial distribution. Broom and Wrigley (1983) and Wrigley and Dunn (1985) have proposed to incorporate explanatory variables into negative binomial and Dirichlet models using loglinear forms. Such extensions clearly are an important step forward in building policy-relevant models of dynamic choice behaviour.

3. MODELS OF VARIETY-SEEKING BEHAVIOUR

Two types of variety-seeking behaviour may be distinguished: structural variety-seeking behaviour and temporal variety-seeking behaviour. Structural variety is the variety that is present within a set of objects whereas temporal variety is the variety that is implied by a sequence of choices (Pessemier, 1985). In the context of spatial choice behaviour temporal variety-seeking is perhaps most important. However, since the basic ideas underlying the models directed at these two types of variety are very similar, we will discuss both approaches in this section.

An initial distinction will be made between two types of models: inventory-based models and non-inventory-based models. The specification of the former type of model is explicitly based on the assumption that the attributes of the chosen alternatives are accumulated in attribute inventories, whereas the latter type of model is not explicitly parameterized in this respect.

3.1. *Inventory-based Variety-seeking Models*

One of the first models of variety-seeking behaviour, originally developed for the case of structural variety, was proposed by McAlister (1979). This so-called model of attribute satiation was built on two basic assumptions:

- attributes are cumulative, implying that the total amount of a particular attribute inherent in a group can be calculated by summing the attribute values across the alternatives belonging to that group;
- the marginal utility for each attribute is a decreasing function.

McAlister selected a quadratic utility function to represent the second assumption. More specifically, it was assumed that the square of the difference between the summed attribute values and an individual's ideal point is an appropriate functional form. Hence, it was assumed that a combination of alternatives will be chosen if:

$$U(g) > U(h) \qquad \forall\ h \neq g \tag{11}$$

where
$$U(g) = - \sum_{k=1}^{K} w_k (X_{g.k} - \hat{X}_k)^2 \tag{12}$$

$U(g)$ is the utility of group g of choice alternatives;
$X_{g.k}$ is the sum of the values for attribute k across the choice alternatives in group g;
$\hat{X}_k$ is the ideal (most preferred) level for the k-th attribute;
w_k is the importance weight of the k-th attribute;
K is the total number of attributes.

In the same study, a more sophisticated version based on Farquhar and Rao's balance model (Farquhar and Rao, 1976) was tested. They divided the attributes in four types. First, they made a distinction between desirable and undesirable attributes. Their relationship to preference is assumed to be reflected in a linearly increasing respectively decreasing function. Second, they assume that preference for a collection of choice alternatives is also influenced by the diversity within the collection. If diversity increases preference, the attribute is called 'counterbalancing'. If, in contrast, preference decreases with increasing diversity, the attribute is termed 'equibalancing'. For both these types of attributes, Farquhar and Rao posit a linear relationship with preference for the collection of choice alternatives.

In a follow-up paper, McAlister (1982) extended the attribute satiation model to the case of temporal variety-seeking. This dynamic attribute satiation model (DAS) differs from its predecessor in that a time-related additional assumption is built in. More specifically it is assumed that a consumption history may be converted into an inventory by a time-related inventory retention factor λ_k. The model may be expressed as follows:

$$U_i^{(T)} > U_j^{(T)} \tag{13}$$

The utility of alternative i at time T is defined as:

$$U_i^{(T)} = - \sum_{k=1}^{K} w_k \left[(I_k^{(T)} + X_{ik}) - \hat{X}_k \right]^2 \tag{14}$$

where $I_k^{(T)}$ is the inventory of attribute k at time T;
X_{ik} is the amount of alternative i on attribute k.

The inventory of attribute k at time T is defined as:

$$I_k^{(T)} = \sum_{t=1}^{T-1} \lambda_k^{T-t} \; X_k^{(t)} \qquad (15)$$

where $X_k^{(t)}$ is the amount of the alternative chosen at time t on attribute k;

λ_k is an inventory retention factor for the k-th attribute, $0 \leq \lambda_k \leq 1$.

Note that if λ_k is 1 for all k, the above model is equal to the attribute satiation model. If $\lambda_k \neq 1$, a consumption history is converted into an inventory by an inverse function of the inventory retention factor.

In order to calibrate these models, the researcher should collect the following data. Subjects' perceptions of the degree in which a choice alternative possesses the selected attributes should be measured. In addition, subjects should be asked to rank the (combinations of) choice alternatives from most likely to least likely to choose at a particular point in time. In case of the DAS model, individuals' consumption histories should also be collected. The ranking task may then be exploited to derive paired comparisons. The parameters of the model, the ideal point and importance weight for each attribute, may then be derived by using linear programming techniques. In case of the DAS model, the inventory retention factors should preferably be calibrated as well, but McAlister assumed these factors all equalled 0.5.

McAlister and Pessemier (1982) extended the DAS model by a term which represents the stimulation contribution to preference. This additional term may be expressed as:

$$w_{K+1}(D_i^{(T)} - \hat{X}_{k+1})^2 \qquad (16)$$

where w_{K+1} represents the importance of the stimulation contribution to preference;

$D_i^{(T)}$ is the total stimulation that will result from choosing alternative i at time T;

$\hat{X}_{K+1}$ is the ideal point for stimulation.

The total amount of stimulation is recursively calculated as:

$$D_i^{(T)} = \lambda_{K+1} D_i^{(T-1)} + \sum_{t=1}^{T-1} \lambda_{K+1}^{T-t} \left[\sum_{k=1}^{K} w_k (X_{ik} - X_k^{(t)})^2 \right] \qquad (17)$$

where λ_{K+1} is a stimulation retention factor;

$D_i^{(0)} = 0.$

Pessemier (1985) has presented a still more sophisticated model of variety-seeking behaviour. He assumed that change in utility results from each attribute of a choice alternative and from interpersonal and intrapersonal variety which the object conveys. Especially, the notion of interpersonal variety is new and it represents an individual's need for group affiliation and personal identity. The model can be best appreciated if it is worked through backwards. An individual's utility for a choice alternative is assumed to be a linear function of the squared distance between the individual's ideal point and the inventory position of that choice alternative in a space of K+2 dimensions. In formula:

$$U_i^{(T)} = a + b\, d_{i,T+1}^2 \tag{18}$$

where $$d_{i,T}^2 = \sum_{k=1}^{K+2} w_k (I_{ik}^{(T)} - \hat{X}_k)^2 \tag{19}$$

$I_{ik}^{(T)}$ is the inventory of the k-th attribute of choice alternative i at time T;

$\hat{X}_k$ is an individual's ideal point for the k-th attribute;

w_k is the importance or salience of the k-th attribute;

a,b are regression coefficients.

The space can be divided into K dimensions associated with the attributes, 1 dimension associated with intrapersonal variety and 1 dimension with interpersonal variety. The definition of the inventories depends on the three kinds of dimensions. The individual inventory level maintained for a particular attribute is assumed to be a function of the times at which increments of the attributes were required, the size of the increments and the consumption rate. The inventory level at time T for attribute k is defined as:

$$I_{ik}^{(T)} = \alpha_k \sum_{t=1}^{T-1} \left[X_k^{(t)} (1 + r_k)^{-t}\right] \qquad k=1,2,\ldots,K \tag{20}$$

where $X_k^{(t)}$ is the amount of attribute k for the choice alternative chosen at time t;

r_k is a time discount rate for attribute k ($r_k \leq 0$);

α_k is a scaling factor for attribute k.

All the r_k's and α_k's are determined by nonlinear least squares methods. The individual ideal points are the dependent observations.

The inventory level of intrapersonal varied experiences is defined by:

$$I_{i,K+1}^{(T)} = \alpha_{K+1} \sum_{t=1}^{T-1} \left[d^*_{t,t-1} (1 + r_{K+1})^{-t} \right] \tag{21}$$

d^* measures the dissimilarity of the alternative chosen at time t to the alternative chosen at t-1. It is defined as the Euclidean distance from alternative i to alternative j, modified by counting only noticeable differences:

$$d^*_{ij} = \left[\sum_{k=1}^{K} X_{ijk}^2 \right]^{0.5} \tag{22}$$

$$\text{where } X_{ijk} = \begin{cases} |X_{ik} - X_{jk}| \text{ if } |X_{ik} - X_{jk}| > c_k X_{min,k} \\ 0. \text{ otherwise} \end{cases} \tag{23}$$

c_k is a simple fraction.

Finally, the interpersonal inventory level is defined by:

$$I_{i,K+2}^{(T)} = \alpha_{K+2} \sum_{t=1}^{T-1} \left[\delta_t (1 + r_{K+2})^{-t} \right] + \alpha_{K+3} \sum_{t=1}^{T-1} \left[\beta_t (1 + r_{K+3})^{-t} \right] \tag{24}$$

where δ_t measures the similarity of an individual ideal point to the mean of the ideal points for friends $(\bar{X}_k^{(t)})$, associates and role models;
β_t measures the individuality of each choice.

More specifically:

$$\delta_t = \left[\sum_{k=1}^{K} (\bar{X}_k{}^{(t)} - X_k{}^{(t)})^2 \right]^{0.5} \tag{25}$$

$$\beta_t = -\ell n \gamma_t \tag{26}$$

where γ_t is the share of valued peer's choices going to the alternative chosen by the individual during the period ending at t.

The ideal points of the attributes are derived from a joint space analysis. Object ratings or paired similarity ratings are used to develop the object space.

3.2. *Non-inventory-based Variety-seeking Models*

The models described in the previous section all are explicitly based in some way on the attributes of the choice alternatives. The models to be discussed in this section are not. Most of these models are based on the concept of first-order Markov chains. In fact, they attempt to predict switching probabilities from concepts of variety-seeking.

A somewhat different model has been suggested by Jeuland (1978). He assumed that the utility of a given choice alternative is a function of the past experience of an individual with that alternative and its unique characteristics. Thus:

$$U_i^{(T)} = f(u_i,\ E_i^{(T)}) \tag{27}$$

where $E_i^{(T)}$ represents the amount of experience with choice alternative i at time T;
u_i accounts for the unique characteristics of choice alternative i.

He assumed that a choice alternative will be chosen if its utility exceeds that of other alternatives by at least a positive constant or threshold Δ; that is:

$$U_i^{(T)} > U_j^{(T)} + \Delta, \qquad \forall\ j \neq i \tag{28}$$

Jeuland postulated the following experience function:

$$E_i^{(T)} = E_i^{(o)}\ \exp(-\lambda T) + \delta_i^{(T-1)} \tag{29}$$

where $E_i^{(o)}$ is the amount of experience with alternative i at the previous time it was chosen;
λ is a parameter which accounts for the declining over time of the experience function in the absence of choices of the alternative;
$\delta_i^{(T)}$ is equal to zero if a choice alternative other than i is chosen at time T, else $\delta_i^{(T)}$ is equal to 1.

The utility expression $U_i^{(T)}$ itself was defined as:

$$U_i^{(T)} = u_i \;/\; [1 + \theta E_i^{(T)}] \tag{30}$$

where θ is a parameter.

Unfortunately, Jeuland only provided simulation results for this model. It was not estimated on real-world data, hence the predictive properties of this model remain unclear.

Givon (1984) proposed a first-order Markov model which is based on the assumption that a variety-seeker evaluates change positively regardless of the alternative previously chosen. His model assumes that the probability of choosing alternative i given that alternative j was chosen on a previous occasion is a function of the preference for choice alternative i and of preference for switching. Individuals may have a negative switching preference. Givon also extended this model to the situation in which an individual partitions choice alternatives according to some underlying attribute and seeks variety by switching among partitions. In this case, the probability of choosing alternative i given that alternative j was chosen on the previous occasion is a function of preference for alternative i and of preference for all alternatives in the partition with alternative j.

This model was further extended by McAlister (1984) and Lattin and McAlister (1985). Following Tversky's (1977) ideas on similarity, they assumed that similarity between choice alternatives is a function of the features they share. The probability $p_{i|j}$ of choosing alternative i given that alternative j was chosen last time then equals:

$$p_{i|j} = \frac{\Pi_i - vS_{ij}}{1 - (v \sum_{i'=1} S_{i'j})} \tag{31}$$

where Π_i is a parameter which reflects the sum of features except the universal ones of alternative i ($\Pi_i \geq 0$, $\sum_i \Pi_i = 1$);

S_{ij} is a parameter which reflects the features shared by alternatives i and j ($0 \leq S_{ij} \leq \min(\Pi_i, \Pi_j)$);

v is a variety-intensity parameter ($0 \leq v \leq 1$).

The model is estimated by solving a constrained optimization problem, which minimizes the sum of squared differences between the observed switching probabilities and the predicted probabilities. More specifically, the estimation task can be written as:

$$\text{Minimize} \sum_i \sum_j \left[p_{i|j} - \frac{\Pi_i - vS_{ij}}{1 - (v \sum_{i'=1} S_{i'j})} \right]^2 \tag{32}$$

4. DYNAMIC DISCRETE CHOICE MODELS

Over the past few years, standard discrete choice theory as applied to static choice behaviour has been extended to the case of dynamic choice behaviour. Seminal work in this area has been conducted by Heckmann (1981), whose major concern was to distinguish between true state dependence and spurious state dependence, resulting from serial correlation. He developed a general framework for analyzing dynamic choice. He assumes a random sample for which data on the presence or absence of an event in each of T equispaced time intervals exists. He also assumes that the event occurs in period t for individual i if and only if a continuous latent variable $Y_{i,t}$ crosses a threshold. This random variable is supposed to consist of two components: a function of exogenous, predetermined and measured endogenous variables that affect current choices and a stochastic disturbance component. The disturbance component may take on various specifications, but in the present paper Heckmann assumed that the disturbances are jointly normally distributed, similar to the multinomial probit model.

Given these assumptions, the general model of dynamic choice may be written as:

$$Y_i^{(t)} = \mathbf{X}_i^{(t)}\boldsymbol{\beta} + \sum_{j=1}^{\infty} \gamma_{t-j,t} d_{i,t-j} + \sum_{j=1}^{\infty} \lambda_{j,t-j} \prod_{\ell=1}^{j} d_{i,t-\ell} + G(L)Y_i^{(t)} + \varepsilon_i^{(t)} \tag{33}$$

where $G(L)$ is a general lag operator of order K, $(G(0) = 0)$;

$d_{i,t}$ is a dummy representing whether the event has occurred $(Y_i^{(t)} \geq 0$ or $Y_i^{(t)} < 0)$;

$\mathbf{X}_i^{(t)}$ represents a set of exogenous variables;

$\boldsymbol{\beta}$ is a parameter vector;

γ and λ are parameters;

$\varepsilon_i^{(t)}$ is an error term.

The first term on the right-hand side of the equation represents the effects of exogenous variables on utilities at time t. The second term represents the effect of the entire past history on choice behaviour at time t. The third term represents the cumulative effect of the most recent continuous experience in a state and the fourth term captures the effect of habit persistence. Heckmann then shows that several models such as Bernoulli models, models with structural state dependence, renewal models, models with general correlations in the error term and habit persistence models can be accommodated in this general model.

Another important publication stems from Tardiff (1979). He suggested the following utility function for the dynamic case:

$$U_{qi}^{(t)} = \mathbf{X}_{qi}^{\prime(t)}\boldsymbol{\beta} + \sum_j \beta^*_{ij} C_{jq}^{(t-1)} + \tilde{\varepsilon}_{qi} + \varepsilon^{*(t)}_{qi} \tag{34}$$

where $U_{qi}^{(t)}$ is the utility of choice alternative i to individual q at time t;

$C_{jq}^{(t-1)}$ is 1 if individual q chooses alternative j at time $(t-1)$ and 0 otherwise;

β and β^* are parameters;

$\tilde{\varepsilon}_{qi}$ is an error term that varies among individuals but not among time periods;

$\varepsilon^{*(t)}_{qi}$ is an error term that varies among both individuals and time periods.

By setting various components of the above utility function at zero, some special cases arise. For example in the cases that $\beta^*_{ij} = 0 \ \forall\ i,j$ and $\varepsilon^{*(t)}_{qi} = 0 \ \forall\ q,i$, the usual type of discrete choice models can be applied directly to the dynamic problem. However, in the case in which the error terms are assumed to be correlated over time, standard estimation procedures are no longer valid. In this case one should either adopt a fixed effects approach, in which the $\tilde{\varepsilon}_{qi}$-terms are explicitly identified and standard discrete choice models are applied directly, or a random effects approach, in which the error variance structure is dealt with directly.

Daganzo and Sheffi (1982) showed that the choice of a state dependence model, a serial correlation model or a hybrid thereof is simply a specification issue, implying that existing computer codes can be used to estimate such models of dynamic choice behaviour. The choice model is specified as:

$$U_i^{(t)} = \boldsymbol{\beta}^{\prime(t)}\ \mathbf{X}_i^{(t)} \tag{35}$$

where $U_i^{(t)}$ is the utility of alternative i at time t;

$\mathbf{X}_i^{(t)}$ is a vector of attribute values for alternative i at time t;

$\boldsymbol{\beta}^{(t)}$ is a vector of parameters at time t.

The vector $\boldsymbol{\beta}'$, $\boldsymbol{\beta}' = (\boldsymbol{\beta}^{(1)},\ldots,\boldsymbol{\beta}^{(t)},\ldots,\boldsymbol{\beta}^{(T)})$ is assumed to be multivariate normal distributed:

$$\boldsymbol{\beta}' \sim \mathrm{MVN}(\bar{\boldsymbol{\beta}}', \Sigma_\beta) \tag{36}$$

If we let c_t denote the choice in period t, the probability of a particular sequence of choices, given conventional discrete choice theory is equal to:

$$p(c_1,\ldots,c_T) = \Pr\left\{U_{c_1}^{(1)} > U_j^{(1)}, \quad \forall\, j \neq c_1; \right. \\ \left. \text{and} \ldots \text{and } U_{cT}^{(T)} > U_j^{(T)}, \quad \forall\, j \neq c_T\right\} \tag{37}$$

Daganzo and Sheffi introduce an auxiliary model which reduces the number of alternatives from N^T, where N is the total number of choice alternatives, to $((N\text{-}1)T + 1)$. In particular, they define a $N(N\text{-}1)$ matrix $\mathbf{\Delta}_i$ such that:

$$\mathbf{X}^{(t)}\mathbf{\Delta}_i = \left[\mathbf{X}_1^{(t)}\text{-}\mathbf{X}_i^{(t)},\ldots,\mathbf{X}_{i-1}^{(t)}\text{-}\mathbf{X}_i^{(t)},\ldots, \right. \\ \left. \mathbf{X}_{i+1}^{(t)}\text{-}\mathbf{X}_i^{(t)},\ldots,\mathbf{X}_N^{(t)}\text{-}\mathbf{X}_i^{(t)}\right] \quad \forall\, i,t \tag{38}$$

implying that the probability equation may be rewritten as:

$$p(c_1,\ldots,c_T) = \Pr\{\boldsymbol{\beta}'\mathbf{X}\mathbf{\Delta} \leq (0,\ldots,0)\} \tag{39}$$

where $\mathbf{X}$ is a $(KT{\times}NT)$ block-diagonal data-attribute matrix for K attributes;
$\mathbf{\Delta}$ is a $(NT{\times}(N\text{-}1)T)$ block diagonal matrix.

This equation corresponds to a multinomial probit function for a $((N\text{-}1)T{+}1)$-dimensional probit model with the following specification:

$$u_0 = 0, \quad (u_1,u_2,\ldots,u(N\text{-}1)T) = \boldsymbol{\beta}'\mathbf{X}\mathbf{\Delta} \tag{40}$$

For the state dependence problem, Daganzo and Sheffi assume the following autoregressive process on the utilities for each time period:

$$\mathbf{U}'^{(t)} = \rho\mathbf{U}'^{(t-1)} + \boldsymbol{\beta}'^{(t)}\mathbf{X}^{(t)} \tag{41}$$

$$\mathbf{U}'^{(t-1)} = \rho\mathbf{U}'^{(t-2)} + \boldsymbol{\beta}'^{(t-1)}\mathbf{X}^{(t-1)} \tag{42}$$

$$\vdots$$

$$\mathbf{U}'^{(1)} = \boldsymbol{\beta}'^{(1)}\mathbf{X}^{(1)} \tag{43}$$

To estimate this model, matrix $\mathbf{X}$ should have the following specification:

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 & \rho\mathbf{X}^{(1)} & \rho^2\mathbf{X}^{(1)} & \dots & \rho^{T-1}\mathbf{X}^{(1)} \\ 0 & \mathbf{X}^{(2)} & \rho\mathbf{X}^{(2)} & \dots & \rho^{T-2}\mathbf{X}^{(2)} \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & \mathbf{X}^{(T)} \end{bmatrix} \tag{44}$$

An application of their approach to two-period panel data can also be found in Johnson and Hensher (1982).

Another dynamic disaggregate choice model has been advanced by Krishnan and Beckmann (1979). Basically their dynamic model is an extension of the Krishnan's static logit model for binary choices which incorporates threshold effects, δ. Introducing time, the basic model can be written as:

$$p(c_t=i) = p(Y^{(t)}=i) + p(c_t=i|Y^{(t)}=3)\ p(Y^{(t)}=3),\ i=1,2 \tag{45}$$

where $p(c_t=i)$ denotes the probability that choice alternative i is chosen at time t;

$p(Y^{(t)}=i)$ is the probability that choice alternative i is preferred at time t;

$p(Y^{(t)}=3)$ is the probability that the individual is indifferent.

The probabilities $p(Y^{(t)}=1)$ and $p(Y^{(t)}=2)$ are defined as:

$$p(Y^{(t)}=1) = 1\ /\ [1 + \exp(V_2^{(t)} - V_1^{(t)} + \delta_{12})] \tag{46}$$

$$p(Y^{(t)}=2) = 1\ /\ [1 + \exp(V_1^{(t)} - V_2^{(t)} + \delta_{21})] \tag{47}$$

where $V_i^{(t)}$ is the deterministic utility component of choice alternative i at time t, $i=1,2$;

δ_{12} and δ_{21} are threshold values.

This implies that an optimal model may be derived by defining the probability that the first alternative will be chosen, given that the individual is indifferent. The authors postulate two models for the indifference state. First, they postulate that if alternative 1 and alternative 2 are equally preferred, alternative 1 will be chosen with probability θ, and alternative 2 will be chosen with probability $(1-\theta)$. Thus, it follows that:

$$p(c_t=1) = p(Y^{(t)}=1) + \theta p(Y^{(t)}=3)$$
$$= \theta\left\{\exp(V_1^{(t)}-V_2^{(t)}+\delta_{21}) \;/\; [1+\exp(V_1^{(t)}-V_2^{(t)}+\delta_{21}]\right\} +$$
$$(1-\theta)\left\{1 \;/\; [1+\exp(V_2^{(t)}-V_1^{(t)}-\delta_{12})]\right\} \quad (48)$$

The second model is based on the indifference postulate that the most recently chosen alternative will also be chosen at time t. Hence we have:

$$\begin{aligned} p(c_t=1|c_{t-1}=2) &= p(Y^{(t)}=1) \\ &= 1 \;/\; [1+\exp(V_2^{(t)}-V_1^{(t)}+\delta_{12})] \end{aligned} \quad (49)$$

and

$$\begin{aligned} p(c_t=1|c_{t-1}=1) &= p(Y^{(t)}=1) + p(Y^{(t)}=3) \\ &= \exp(V_1^{(t)}-V_2^{(t)}+\delta_{21}) \;/\; [1+\exp(V_1^{(t)}+\delta_{21})] \end{aligned} \quad (50)$$

Another interesting development is the beta logistic model (Heckmann and Willis, 1977). This model provides predictions for both the mean probability of choosing a particular alternative and the distribution of the choice probabilities around the mean. The original model is based on the assumptions that the exogenous variables are constant over time and the absence of state/time dependence. Heterogeneity is introduced by defining a subgroup (g) of a sample in which all individuals have exactly the same values on all exogenous variables included in the model. The distribution of the probabilities for such a group represents heterogeneity. Given these assumptions the beta logistic model assumes that the mean probability can be modelled in terms of a conventional logit model, and that the distribution of probabilities is a beta distribution. The dichotomous model may be written as:

$$E(p_g) = \frac{\exp\{\mathbf{X}_g' \, (\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2)\}}{1 + \exp\{\mathbf{X}_g' \, (\boldsymbol{\beta}_1 - \boldsymbol{\beta}_2)\}} \quad (51)$$

where $E(p_g)$ is the mean probability of subgroup g choosing an alternative;
$\mathbf{X}_g$ is a set of exogenous variables for subgroup g;
$\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ are vectors of parameter values.

The form of the heterogeneity is given by:

$$f(p_g|\alpha_{1g},\alpha_{2g}) = \frac{\Gamma(\alpha_{1g} + \alpha_{2g})}{\Gamma(\alpha_{1g})\ \Gamma(\alpha_{2g})}\ p_g^{\alpha_{1g}-1}\ (1-p_g)^{\alpha_{2g}-1} \tag{52}$$

where $\alpha_{1g} = \exp\{\mathbf{X}_g'\ \boldsymbol{\beta}_1\}$ (53)

$\alpha_{2g} = \exp\{\mathbf{X}_g'\ \boldsymbol{\beta}_2\}$ (54)

The original specification of the dichotomous beta logistic model has been generalized by Davies (1984) and Davies and Pickles (1984) to incorporate feedback effects and time-varying exogenous variables. They applied the model successfully to the study of residential mobility. The dichotomous model may be extended to the polytomous case by replacing the binomial distribution of the logistic model by the multinomial distribution and the beta distribution describing heterogeneity by the Dirichlet distribution, its multivariate equivalent. Dunn and Wrigley (1985) provide an application of this model in a study of spatial shopping behaviour.

The above models of dynamic choice behaviour can all be considered as extensions of the econometric methodology associated with discrete choice models when using panel data. Leonardi (1983) approached the problem from a different angle and developed a theoretical model. In particular, he assumed that a utility or disutility is associated both with a transition from one alternative to another and with staying with a particular alternative. In addition he assumed a discount rate of utilities over time. It follows that the utility associated with a transition at time $t + \Delta$ is given by:

$$U_j^{(t+\Delta)} = (1 - \alpha\Delta)\ [v_{ij} + V_j^{(t+\Delta)} + \varepsilon_j] \tag{55}$$

where α is the discount rate, $(\alpha > 0)$;

v_{ij} is the utility (disutility) associated with a transition from i to j;

$V_j^{(t+\Delta)}$ is the total expected utility for a process started in i at time $t+\Delta$;

ε_j is an error term.

Likewise, the gain in utility of remaining in i starting from $t + \Delta$ equals:

$$U_i^{(t+\Delta)} = (1 - \alpha\Delta)\ [V_i^{(t+\Delta)} + \varepsilon_i] \tag{56}$$

and he will gain $h_i^{(t)}\Delta$, during his stay in i in $(t,t+\Delta)$, assuming decisions to move are made at the end of the time interval. Assuming utility-maximizing behaviour and independently and identically Gumbel distributed error terms, Leonardi shows that the probability of moving from i to j in $(t,t+\Delta)$ is equal to:

$$p_{ij}^{(t,t+\Delta)} = \frac{M_j \exp\{\beta\ [v_{ij} + V_j^{(t+\Delta)}]\}}{\sum_{j'} M_{j'} \exp\{\beta\ [v_{ij'} + V_{j'}^{(t+\Delta)} + \exp[\beta V_i^{(t+\Delta)}]\}} \qquad (57)$$

where M_j is the number of sampled alternatives.

Leonardi also demonstrated that this choice process can also be formulated in terms of an optimal control problem.

Another interesting theoretical model has been put forward by De Palma. Basically he has been concerned with an expansion of classical models of individual behaviour by taking into account interindividual interaction and interdependence of individual decisions (Deneubourg, De Palma and Lefèvre 1979; De Palma and Lefèvre, 1981, 1983; De Palma, 1983). The deterministic utility component consists of two parts: a part representing the absolute benefit associated with a choice and a part which measures the relative benefit associated with a choice given the choice behaviour of another actor. By assuming that each actor anticipates rationally the behaviour of the other actor and linear utility functions, De Palma derives a theoretical model which is able to describe such phenomena as repulsion, attraction and competition between actors. The model may be viewed as dynamic in the sense that time could be used as a basis for the utility function specification. If not, his model should be considered as an extension of conventional static discrete choice models.

The field of dynamic discrete choice models is developing rapidly. Apart from the contributions discussed above, several other important publications have appeared recently. For example, Avery *et al.* (1983) presented a multiperiod probit model; Hensher (1984) developed a quasi-dynamic choice model for automobile demand; Dagsvik (1983) introduced a dynamic extension of Thurstonian and Lucian choice models; Meyer and Sathi (1985) discussed a dynamic model of consumer choice involving product learning; and Miller and O'Kelly (1983) used a dynamic logit model to estimate shopping destination choice. The reader is referred to these publications for further details.

5. CONCLUDING REMARKS

It is evident that the modelling of dynamic aspects of choice behaviour is getting increasingly more attention in geography and related disciplines. It is to be expected therefore that in the next few years most important advances will be made in this area. While much progress has already been made, it is also true that existing models are still based on very rigorous assumptions such as stationarity. In any case, the various effects discussed for conventional discrete choice models still have to be incorporated into their dynamic counterparts. In addition, comparative testing of the models is important to learn their characteristics and predictive performance.

Urban Planning Group
Faculty of Architecture, Building and Planning
Eindhoven University of Technology
The Netherlands

BIBLIOGRAPHY

Avery, R.B., L.P. Hansen, and V.J. Hotz: 1983, 'Multiperiod probit models and orthogonality condition estimation', *International Economic Review* **24**, 21-35.

Broom, D. and N. Wrigley: 1983, 'Incorporating explanatory variables into stochastic panel-data models of urban shopping behavior', *Urban Geography* **4**, 244-257.

Burnett, P.: 1974, 'A three state Markov model of choice behavior within spatial structures', *Geographical Analysis* **6**, 53-68.

Burnett, P.: 1975, 'A Bernoulli model of destination choice', *Transportation Research Record* **527**, 33-44.

Burnett, P.: 1977, 'Tests of a linear learning model of destination choice: applications to shopping travel by heterogeneous groups', *Geografiska Annaler* **B 59**, 95-108.

Burnett, P.: 1978, 'Markovian models of movement within urban spatial structures', *Geographical Analysis* **10**, 142-153.

Chatfield, C., A.S.C. Ehrenberg, and G.J. Goodhardt: 1966, 'Progress on a simplified model of stationary purchasing behaviour', *Journal of the Royal Statistical Society Series* **A 129**, 317-367.

Crouchley, R., R.B. Davies, and A. Pickles: 1982, 'A re-examination of Burnett's study of Markovian models of movement', *Geographical Analysis* **14**, 260-262.

Daganzo, C.F. and Y. Sheffi: 1982, 'Multinomial probit with time series data: unifying state dependence and serial correlation models', *Environment and Planning* **A-14**, 1377-1388.

Dagsvik, J.K.: 1983, 'Discrete dynamic choice: an extension of the choice models of Thurstone and Luce', *Journal of Mathematical Psychology* **27**, 1-43.

Davies, R.B.: 1984, 'A generalized beta-logistic model for longitudinal data', *Environment and Planning* **A 16**, 1375-1386.

Davies, R.B. and A.R. Pickles: 1984, 'Accounting for omitted variables in the analysis of panel and other longitudinal data', Papers in Planning Research 77, University of Wales, Institute of Science and Technology, Cardiff.

De Palma, A.: 1983, 'Incomplete information, expectation, and subsequent decisionmaking', *Environment and Planning* **A 15**, 123-130.

De Palma, A. and C.E. Lefèvre: 1981, 'A probabilistic search model', *Journal of Mathematical Sociology* **8**, 43-60.

De Palma, A. and C.E. Lefèvre: 1983, 'Individual decision making in dynamic collective systems', *Journal of Mathematical Sociology* **9**, 103-124.

Deneubourg, J.L., A. De Palma, and D. Kahn: 1979, 'Dynamic models of competition between transportation modes', *Environment and Planning* **11**, 665-673.

Dunn, R., S. Reader, and N. Wrigley: 1983, 'An investigation of the assumptions of the NBD model as applied to purchasing at individual stores', *Applied Statistics* **32**, 249-259.

Dunn, R. and N. Wrigley: 1985, 'Beta-logistic models of urban shopping center choice', *Geographical Analysis* **17**, 95-113.

Ehrenberg, A.S.C.: 1968, 'The practical meaning and usefulness of the NBD/LSD theory of repeat buying', *Applied Statistics* **17**, 17-32.

Ehrenberg, A.S.C.: 1972, *Repeat Buying: Theory and Application*, North-Holland, Amsterdam.

Farquhar, P.H. and V.R. Rao: 1976, 'A balance model of evaluating subsets of multiattribute items', *Management Science* **22**, 528-539.

Ginsberg, R.B.: 1971, 'Semi-Markov processes and mobility', *Journal of Mathematical Sociology* **1**, 233-263.

Ginsberg, R.B.: 1972, 'Incorporating causal structure and exogenous information within probabilistic models, with special reference to choice, gravity, migration and Markov chains', *Journal of Mathematical Sociology* **2**, 88-103.

Ginsberg, R.B.: 1973, 'Stochastic models of residential and geographic mobility for heterogenous populations', *Environment and Planning* **A 5**, 113-124.

Ginsberg, R.B.: 1978, 'The relationship between timing of moves and choices of destination in stochastic models of migration', *Environment and Planning* **A 10**, 667-679.

Ginsberg, R.B.: 1979, 'Tests of stochastic models of timing in mobility histories', *Environment and Planning* **A 11**, 1387-1404.

Givon, M.: 1984, 'Variety seeking through brand switching', *Marketing Science* **3**, 1-22.

Goodhardt, G.J., A.S.C. Ehrenberg, and C. Chatfield: 1984, 'The Dirichlet: a comprehensive model of buying behaviour', *Journal of the Royal Statistical Society Series* **A 147**, 621-655.

Halperin, W.C.: 1985, 'The analysis of panel data for discrete choices', in P. Nijkamp, H. Leitner and N. Wrigley (eds.), *Measuring the Unmeasurable: the Analysis of Qualitative Spatial Data*, Martinus Nijhoff, The Hague.

Halperin, W.C. and N. Gale: 1984, 'Towards behavioural models of spatial choice: some recent developments', in D.E. Pitfield (ed.), *Discrete Choice Models in Regional Science*, Pion, London.

Heckman, J.J.: 1981, 'Statistical models for discrete panel data', in C.F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, Cambridge, Massachusetts.

Heckman, J.J. and R. Willis: 1977, 'A beta-logistic model for the analysis of sequential labor force participation by married women', *Journal of Political Economy* **85**, 27-58.

Hensher, D.A.: 1984, 'Model specification for a quasi-dynamic discrete-continuous choice automobile demand system in discrete time using panel data', Working Paper No 11, School of Economic and Financial Studies, Marquarie University, North Ryde.

Hensher, D.A. and N. Wrigley: 1984, 'Statistical modelling of discrete choices with panel data', Working Paper No 16, Dimensions of Automobile Demand Project, School of Economic and Financial Studies, Marquarie University, North Ryde.
Herniter, J.: 1971, 'A probabilistic market model of purchase timing and brand selection', *Management Science* **18**, 102-113.
Jeuland, A.: 1978, 'Brand preference over time: a partially deterministic operationalization of the notion of variety seeking', in S. Jain (ed.), *Research Frontiers in Marketing: Dialogues and Directions*, American Marketing Association, Chicago.
Jeuland, A.B., F.M. Bass, and G.P. Wright: 1980, 'A multibrand stochastic model compounding heterogenous Erlang timing and multinomial choice processes', *Operations Research* **28**, 255-277.
Johnson, L. and D.A. Hensher: 1982, 'Application of multinomial probit to a two-period panel data set', *Transportation Research* A **16**, 457-464.
Jones, J.M. and F.S. Zufryden: 1980, 'Adding explanatory variables to a consumer purchase behavior model: an exploratory study', *Journal of Marketing Research* **17**, 323-334.
Jones, J.M. and F.S. Zufryden: 1982, 'An approach for assessing demographic and price influences on brand purchase behavior', *Journal of Marketing* **46**, 36-46.
Kau, K.A. and A.S.C. Ehrenberg: 1984, 'Patterns of store choice', *Journal of Marketing Research* **21**, 399-409.
Krishnan, K.S. and M.J. Beckmann: 1979, 'Dynamic disaggregate choice models with an application in transportation', *Decision Science* **10**, 218-231.
Lattin, J.H. and L. McAlister: 1985, 'Using a variety seeking model to identify substitute and complementary relationships among competing products', *Journal of Marketing Research* **9**, 141-151.
Leonardi, G.: 1983, 'An optimal control representation of a stochastic multistage multiactor choice process', in D.A. Griffith and A.C. Lea (eds.), *Evolving Geographical Structures: Mathematical Models and Theories for Space-Time Processes*, Martinus Nijhoff, The Hague.
Massy, W.F., D.B. Montgomery, and D.G. Morrison: 1970, *Stochastic models of buying behaviour*, MIT Press, Cambridge.
McAlister, L.: 1979, 'Choosing multiple items from a product class', *Journal of Consumer Research* **6**, 213-224.
McAlister, L.: 1982, 'A dynamic attribute satiation model of variety seeking behavior', *Journal of Consumer Research* **9**, 141-151.
McAlister, L.: 1984, 'A similarity based Markov model of variety seeking behavior, its implications and a linear programming technique to estimate its parameters', MIT Working Paper 1539-84, A.P. Sloan School of Management, Cambridge, Massachusetts.
McAlister, L. and E.A. Pessemier: 1982, 'Variety-seeking behavior: an interdisciplinary review', *Journal of Consumer Research* **9**, 311-322.
Meyer, R.J. and A. Sathi: 1985, 'A multiattribute model of consumer choice during product learning', *Marketing Science* **4**, 41-61.
Miller, E.J. and M.E. O'Kelly: 1983, 'Estimating shopping destination models from travel diary data', *The Professional Geographer* **35**, 440-449.

Paull, A.E.: 1978, 'A generalized compound Poisson model for consumer purchase panel data analysis', *Journal of the American Statistical Association* **73**, 706-713.

Pessemier, E.A.: 1985, 'Varied individual behavior: some theories, measurement and models', *Multivariate Behavioral Research* **20**, 69-94.

Tardiff, T.J.: 1979, 'Definition of alternatives and representation of dynamic behaviour in spatial choice models', *Transportation Research Record* **723**, 25-30.

Tversky, A.: 1977, 'Features of similarity', *Psychological Review* **84**, 327-352.

Wrigley, N.: 1980, 'An approach to the modelling of shop-choice patterns: an explanatory analysis of purchasing patterns in a British city', in D.T. Herbert and R.J. Johnston (eds.), *Geography and the Urban Environment: Progress in Research and Applications*, Volume 3, John Wiley, Chichester, Sussex.

Wrigley, N. and R. Dunn: 1984a, 'Stochastic panel-data models of urban shopping behaviour: 1.purchasing at individual stores in a city', *Environment and Planning* **A 16**, 629-650.

Wrigley, N. and R. Dunn: 1984b, 'Stochastic panel-data models of urban shopping behaviour: 2.multistore purchasing patterns and the Dirichlet model', *Environment and Planning* **A 16**, 759-778.

Wrigley, N. and R. Dunn: 1984c, 'Stochastic panel-data models of urban shopping behaviour: 3.the interaction of store choice and brand choice', *Environment and Planning* **A 16**, 1221-1236.

Wrigley, N. and R. Dunn: 1985, 'Stochastic panel-data models of urban shopping behaviour: 4.incorporating independent variables into the NBD and Dirichlet models', *Environment and Planning* **A 17**, 319-332.

Zufryden, F.S.: 1977, 'A composite heterogeneous model of brand choice and purchase timing behavior', *Management Science* **24**, 121-136.

Zufryden, F.S.: 1978, 'An empirical evaluation of a composite heterogeneous model of brand choice and purchase timing behavior', *Management Science* **24**, 761-773.