

# Random utility models in marketing research: a survey

George Baltas<sup>a,\*</sup>, Peter Doyle<sup>b</sup>

<sup>a</sup>*Athens University of Economics and Business, Athens, Greece*

<sup>b</sup>*Warwick Business School, University of Warwick, Coventry CV4 7AL, England, UK*

Accepted 1 February 2000

---

## Abstract

Random utility (RU) models are well-established methods for describing discrete choice behavior. Recently, there has been a strong upsurge in interest driven by advances in data gathering and estimation technology. This review paper describes the principles and issues, and develops a taxonomy of three major families of models. The paper summarizes and classifies the different approaches. The advantages and limitations of the various alternatives are outlined. Practical issues in implementing the models are also discussed. © 2000 Elsevier Science Inc. All rights reserved.

*Keywords:* Random utility models; Marketing research; Discrete choice behavior

---

## 1. Introduction

With an ever increasing importance of market intelligence, the need to understand advanced methods of market research has never been greater. Random utility (RU) models have been developed to describe choice among mutually exclusive discrete alternatives and received considerable academic and industry attention. This paper surveys RU models, discusses key issues and develops a structurally meaningful synthesis of the different formulations. It is intended to help managers and researchers keep informed of a fast-changing and important area which may not directly fall within their own specific professional or research interests. The paper is organized as follows. The subsequent section illustrates the underlying principles of RU models and explains how probabilistic choice flows from utility maximization. The third section discusses the behavior of RU models and issues that arise from restrictive stochastic assumptions. The fourth section introduces a taxonomy and describes the different models. The fifth section is concerned with implementation and experimental data. The sixth section discusses some practical issues and the last section concludes.

## 2. Overview

Consider an individual agent choosing a single option among a finite set of alternatives, for example, a consumer deciding which brand to buy. This is the realm of behavior that is considered in RU modeling. In RU models, preferences for such discrete alternatives are determined by the realization of latent indices of attractiveness, called product utilities. Utility maximization is the objective of the decision process and leads to observed choice in the sense that the consumer chooses the alternative for which utility is maximal. Individual preferences depend on characteristics of the alternatives and the tastes of the consumer. An RU model defines a mapping from observed characteristics into preferences. The analyst however cannot observe all the factors affecting preferences and the latter are treated as random variables. By its abstraction from various idiosyncratic factors, the model uses stochastic assumptions to describe unmeasured variation in preferences. An operational way to allow for maximization of latent preferences is to consider a utility function that is decomposable into two additively separable parts, (1) a deterministic component specified as a function of measured attributes of the alternatives and/or the individual, and (2) a stochastic component representing unobserved attributes affecting choice, interindividual differences in utilities depending upon the heterogeneity in tastes, measurement errors, and functional misspecification (Manski, 1977). In the next sections, we shall consider

---

\* Corresponding author.

*E-mail address:* gb@aub.gr (G. Baltas).

different models based on alternative hypotheses about the “unknown.” Proceeding further in the same vein, let

$$U_{ij} = V_{ij} + \varepsilon_{ij} \tag{1}$$

be the utility of alternative  $j$  for consumer  $i$ , where  $V_{ij}$  is the deterministic component and  $\varepsilon_{ij}$  the random component. Typically, the deterministic component  $V_{ij}$  has been assumed to have an additively separable linear form  $V_{ij} = \mathbf{x}_{ij}^T \boldsymbol{\beta}$ , where  $\mathbf{x}_{ij}$  and  $\boldsymbol{\beta}$  are the vectors of exogenous variables and parameters, respectively. In the hypothetical case that  $V$  contains perfect information about the determinants of utility, the consumer would simply choose the product with the highest  $V_{ij}$ . The stochastic terms  $\varepsilon_{ij}$  shaping the true and latent utility in Eq. (1), introduce uncertainty regarding the choice and therefore, choice probabilities are invoked to describe choice behavior. The probabilistic description of choice has been introduced not to reflect behavior that is probabilistic. Rather, it is the lack of information that leads the analyst to treat utility as a random variable and consequently to describe choice in a probabilistic fashion. In fact, the properties of RU models can be attributed to the specific assumptions that each model implies about the stochastic terms. Under the utility maximization rule, a consumer facing a set of available products  $C = \{1,2,3, \dots, M\}$  will choose a product  $j$  with probability  $P(j) = P(U_j > U_k)$  for all  $k \in C, k \neq j$ , or as it follows from Eq. (1):

$$P(j) = P(\varepsilon_k < V_j - V_k + \varepsilon_j) \text{ for all } k \in C, k \neq j. \tag{2}$$

The probability that  $j$  is chosen is then obtained by making assumptions about the form of the distribution of the random variables and integrating Eq. (2) over a continuum of all possible values for  $\varepsilon_j$ . From Eq. (2), we can write the selection probability for, say, the first alternative as

$$P(1) = \int_{-\infty}^{\infty} \int_{-\infty}^{V_1 - V_2 + \varepsilon_1} \int_{-\infty}^{V_1 - V_3 + \varepsilon_1} \dots \times \int_{-\infty}^{V_1 - V_M + \varepsilon_1} f(\varepsilon_1, \varepsilon_2, \varepsilon_3, \dots, \varepsilon_M) d\varepsilon_M \dots d\varepsilon_3 d\varepsilon_2 d\varepsilon_1. \tag{3}$$

In words, Eq. (3) states that the choice probability of alternative 1 is the probability of over all possible values of  $\varepsilon_1$ , all the other random terms being less than  $V_1 - V_j + \varepsilon_1, \forall j \in C$ . The popular multinomial logit (MNL) model is derived by assuming that the random terms are independently identically distributed (IID) according to the double exponential distribution with mode zero and variance  $\mu^2 \pi^2 / 6$ , where  $\mu$  is a positive scale parameter. The choice probability in Eq. (3) then takes the compact form,

$$P(j) = \frac{\exp\left(\frac{V_j}{\mu}\right)}{\sum_{k \in C} \exp\left(\frac{V_k}{\mu}\right)}. \tag{4}$$

The analytic form of the MNL probabilities has greatly contributed to the popularity of the MNL model. The

expression in Eq. (4) can be derived in a great number of ways (McFadden, 1973; Train, 1986; Anderson et al., 1992). Having laid out the necessary background, we turn to the stochastic assumptions of the models.

### 3. Stochastic assumptions of RU models

Suppose we observe members of a population of consumers, each member  $i$  of which has a utility function  $U_{ij} = V_{ij} + \varepsilon_{ij}$  for each product  $j$  of a set  $C = [1,2,3, \dots, M]$ .  $V_{ij}$  is the non-stochastic function mapping attributes into utility and  $\varepsilon_{ij}$  accounts for factors not included in  $V_{ij}$ .

The simple MNL model accounts for unobserved determinants of choice by IID random terms. That is they are assumed to have the same distribution, with the same mean and variance and also to be uncorrelated across and within individuals. An interesting property is the effect of increasing unexplained stochastic variation on the identified coefficients. Since the variance (assumed the same for all  $\varepsilon_j$ ) is related with the parameter  $\mu$ , it is obvious in Eq. (4) that the variance discounts the value of the estimated parameters in the non-stochastic function  $V$ . Since the variables in  $V$  are exogenous, the estimated coefficients absorb the variance effect. Intuitively, high variance implies limited ability of the observed variables to explain choices and therefore leads to smaller values of the coefficients. Since  $\mu$  is a transformation of the variation of the random disturbances, it can be seen as an index of unobserved variation in preferences that cannot be explained by the variables in the non-stochastic function  $V$ . In the MNL, the price coefficient reflects the response of choice probabilities to prices and its magnitude is related to the variance parameter  $\mu$ . As unobserved variation decreases, the value of the identified price parameter in  $V$  increases and vice versa. Therefore, the identified price coefficient can be regarded as an index of average substitutability among alternatives that is related to stochastic variation. Although the MNL accommodates varying rates of symmetric substitution, the assumption of IID random components remains restrictive and imposes the independence of irrelevant alternatives (IIA) property (Ben-Akiva and Lerman, 1985). Under this structural restriction, the odds of the consumer choosing  $j$  over  $k$  remain the same regardless of the composition of the choice set. An analogous and possibly more important drawback is that the model cannot postulate any pattern of differential substitutability between products. An improvement in an alternative’s systematic utility will have a proportionally equal impact on the selection probabilities of all other alternatives. Thus, an implication of the IIA property is that the cross-elasticity of the probability of brand  $j$  with respect to a change in  $V_k$  is the same for all  $j$  with  $j \neq k$ .

The assumption of independent preferences is restrictive. In reality, alternatives may not be equally dissimilar. Differential similarities among products due to shared characteristics lead to correlated utilities. When these conditions

apply, the random utility components are not independent. An important implication is that the rates of substitution are not uniform since cross-alternative substitutability is related to perceived similarity (McFadden, 1981).

Similarly, the assumption of identically distributed utilities may not be pragmatic. Unmeasured factors may have unequal variances across alternatives. For instance, the suitability on consumption occasions may vary more for brand A than for brand B. Similarly, interpersonal differences in preferences may not be uniform for all brands, that is some brands have utilities that vary more over people. Where these conditions apply, the non-stochastic function  $V$  does not explain equally well choices of different alternatives (Baltas and Doyle, 1998).

When an RU model is used as a model of population behavior, another issue may arise because of parameter heterogeneity across consumers. From a formal point of view, the stochastic utility of alternative  $j$  for individual  $i$  at time  $t$ ,  $\varepsilon_{ijt}$ , embodies both interindividual and intraindividual unmeasured variation in preferences. The interindividual variation may include unobserved factors, such as intrinsic brand preference and deviations from average population sensitivity to marketing variables, which are persistent over time for each individual. The intraindividual variation may include unobserved factors, such as different states of mind, different consumption occasions and dynamic taste formation, which vary over time for each individual.

To appreciate these ideas consider that in the specification  $U_{ijt} = \mathbf{x}_{ijt} \beta + \varepsilon_{ijt}$ , the parameter vector describes mean population tastes,  $E b_i = \beta$ , where  $b_i$  may vary over individuals. The unobserved random component can be expressed as  $\varepsilon_{ijt} = \mathbf{x}_{ijt} (\beta - b_i) + \xi_{ijt}$  where  $\xi_{ijt}$  is a random term that is assumed IID over alternatives and time. Therefore, the unobserved component  $\varepsilon_{ijt}$  is correlated over  $j$  because of the presence of the individual specific taste deviation ( $\beta - b_i$ ) (e.g. Train, 1998).

Unobserved heterogeneity may create correlation not only over alternatives but also over time. Since the realization of the random variable  $U_{ij}$  over successive occasions  $t$  depends on the invariable component ( $\beta - b_i$ ), the utilities attached to a product are serially correlated. Unobserved heterogeneity would appear as serial correlation in the random terms (Keane, 1997b).

Heterogeneity however is not the only reason for serial correlation. Another possible explanation is the presence of intrapersonal dynamics. State dependence and short-run preference inertia can also lead to serially correlated utilities. As the name suggests, state dependence is the influence of past outcomes (states) on present choices. On the other hand, preference inertia or habit-persistence (Heckman, 1981a; Roy et al., 1996) is the dependence of current utilities on past utilities. A main behavioral distinction between heterogeneity and inertia is that heterogeneity refers to time-invariant patterns while inertia describes behavior that tends to be similar in the short-run (Keane,

1997a). Similarly, state dependence refers to effects of prior experience in a clear causal sense while inertia describes short-run persistency. In conclusion, utilities may be correlated over time within people because they are evaluated by the same tastes (heterogeneity), because they are determined by past experiences (state dependence), or because they depend on past utilities (inertia). We shall consider interpersonal heterogeneity and intrapersonal dynamics as two related forms of taste variation, i.e. variation over people and variation over time.

To summarize: (a) unobserved product attributes may lead to non-IID (over products) random utilities, (b) unobserved taste heterogeneity may lead to non-IID (over products) random utilities across individuals and also to non-IID (over time) random utilities within individuals, (c) intrapersonal dynamics may lead to non-IID (over time) random utilities within individuals.

These distinctions are instructive. For example, a homogeneous probit model accounts for unobserved product attributes but does not deal explicitly with interindividual differences. Analogously, as we discuss, subsequently, a random-coefficients MNL model account for interindividual differences but does not deal explicitly with unobserved product attributes. Although both are free from the IIA property when used as models of population behavior, their structural flexibility has different origins. For the first, it arises out of an approximation of individual decision that accounts for unobserved product attributes. For the second, it flows from a description of population behavior that accounts for aggregation over heterogeneous utility functions. We now turn to describe and integrate alternative formulations that lie within the class of RU models.

#### 4. A taxonomy of RU models

We propose a characterization of the models along the following lines. As Fig. 1 illustrates, the major branches represent the three major issues: (a) unobserved heterogeneity in alternatives, (b) taste variation, and (c) heterogeneous choice sets.

The first branch describes alternative treatments of unobserved product attributes. First, we distinguish between models that assume IIA and non-IIA unobserved product heterogeneity. The first class includes only the MNL. In theory, other models (e.g. a restricted probit with IID terms) can be cast as members of the same class. In practice, only the MNL has been used. The second class allows utilities that are non-IID over products and account for unobserved product heterogeneity that violates the IIA property. Consequently, the models can postulate more flexible patterns of substitution. McFadden's (1979) generalized extreme value (GEV) model, Hausman and Wise's (1978) multinomial probit (MNP) model, and the heteroscedastic extreme value (HEV) model introduced by Daganzo (1979) can be regarded as members of this class.

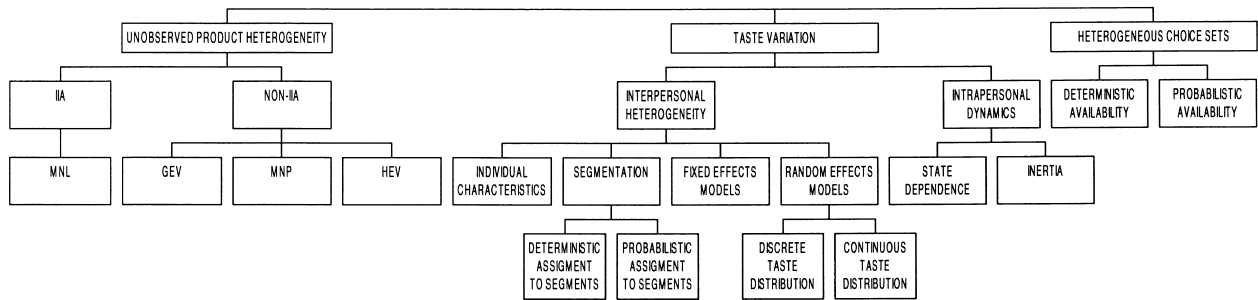


Fig. 1. A taxonomy of random utility models.

The second branch deals with variation in tastes. The first class accounts for taste variation over people, that is, heterogeneous preferences and responses to product characteristics and marketing variables. Several different approaches have been used to account for problems of aggregation over heterogeneous utility functions. One may include individual background variables in the utility function, assign consumers to segments, or allow parameters to vary over people either by fixed or random effects (FE) models (e.g. Chintagunta et al., 1991).

The second class considers taste variation over time. As mentioned earlier, in the presence of choice dynamics, individual tastes change over time in response to previous experiences. However, the determination of dynamic effects in RU models requires explicit control of heterogeneity (see, e.g. Roy et al., 1996). Distinguishing the effect of heterogeneity from the impact of the past on current tastes and consequently, choices is a central issue in empirical work.

To this point, the marketing literature has followed or extended ideas with econometric origins. The third family however can be viewed as a departure from this tradition. This stream of research attempts to include choice set effects in RU models, a concept that has been given much attention in marketing (see Shocker et al., 1991 for a review). The inclusion of heterogeneous choice sets lead to models that describe the idiosyncratic composition of the group of alternatives evaluated prior to choice. As will be seen below, this heterogeneity in the availability of alternatives is created by a multistage process where individuals reduce their options before reaching a final choice.

The following sections discuss the structure and individual models of the taxonomy.

#### 4.1. Unobserved product heterogeneity

The first criterion in classifying RU models is how they deal with unobserved attributes of the choice alternatives. The most simple way is to assume the IIA property. As noted above, this is the assumption of the MNL model. The MNL has played a historical role and it is still

important for applied work. The transparent and elegant probability formula allows a clear interpretation in terms of the relative utilities. In addition, in practice, estimation and forecasting are easy. The share of the new product will equal the proportional decreases in the share of all existing products and the response of demand to policy variables is given by compact own and cross elasticities. The MNL can also be estimated on a subset of alternatives, when the IIA is sustained. However, as is often the case, the unmatched tractability of the MNL comes at a cost in flexibility.

In early marketing studies, the MNL model has been used to study a new product strategy (Hauser and Urban, 1977; Silk and Urban, 1978), choice of business schools (Punj and Staelin, 1978), and store choice (Gensch and Recker, 1979). Guadagni and Little (1983) used scanner panel data to calibrate an MNL model. Since then, scanner data have become the dominant source of observations in applied RU analyses. The MNL model has been used to study several other issues such as variety seeking (Lattin, 1987), reference-dependent behavior (Winer, 1986; Lattin and Bucklin, 1989; Hardie et al., 1993), promotion effects (Allenby and Rossi, 1991; Bronnenberg and Wathieu, 1996), advertising effects (Tellis 1988, Deighton et al., 1994), brand equity (Swait et al., 1993) and price effects (Krishnamurthi and Raj, 1988, 1991).

The models included in the non-IIA class describe unobserved product heterogeneity that violates the IIA property. In practical terms, these models relax the stochastic structure of the MNL model by accommodating preferences that are correlated, non-identical, or both across alternatives (Fig. 1).

The standard GEV or nested logit (NMNL) model (Ben-Akiva, 1973; McFadden, 1979) partitions the set of available alternatives into subsets of relatively homogeneous alternatives and postulates that substitution is greater within than between subsets. This choice structure appears like a tree with branches defining choices among group of elemental alternatives and twigs representing choices of alternatives within branches. The corresponding transition probabilities are MNL. The choice probabilities can be

expressed as products of respective transition probabilities or can be written compactly,

$$P(j, q) = \frac{\exp(V_{jq}/\lambda_q) \left[ \sum_{k \in q} \exp(V_{kq}/\lambda_q) \right]^{(\lambda_q - 1)}}{\sum_{m=1}^n \left[ \sum_{k \in m} \exp(V_{km}/\lambda_m) \right]^{\lambda_m}} \quad (5)$$

where  $V_{jq}$  is the non-stochastic utility for alternative  $j$  in subset  $q$  and  $\lambda_q$  is an index of the inverse of utility correlation or alternative dissimilarity within subsets. The model allows alternatives in the same subset to have correlated utilities as they share unobserved attributes. In addition, the NMNL allows the variance to differ across subsets. Although the NMNL is often used to describe hierarchical choice, it does not require a multistage decision process. The nested structure simply illustrates the correlation among the random utility components. The NMNL is popular in marketing (e.g. Dubin, 1986; Buckley, 1988; Kannan and Wright, 1991; Kamakura et al., 1996; Baltas et al., 1997). Among all the non-IIA models, the NMNL is the most empirically tractable. Estimation, forecasting, and evaluation of market response (elasticities) are straightforward. The current approach however is with limitations. While the IIA property does not hold between groups of choices, it is still imposed within each group. This can be regarded as a restriction of the model. Also, the analyst has to specify a priori the partitioning of the choice set in relatively homogeneous groups of alternatives. This however allows specific hypotheses to be tested and prevents ad hoc rationalization of empirical results.

The MNL probabilities assume a common scale parameter  $\mu$  for all random components  $\varepsilon_i$  implying equal variances across all choice alternatives. The HEV model (see Fig. 1) employs the double exponential distribution for the random terms but allows them not to be identically distributed. More specifically, the assumption of common variance is relaxed by allowing alternative-specific scale parameters  $\mu_j$ . The HEV nests the MNL, i.e. the MNL is a constrained form of the HEV (Bhat, 1995). Estimation of the HEV model yields estimates for the parameter vector of the explanatory variables and also for the scale parameters of the stochastic terms. The HEV accommodates differential competitive effects by controlling the impact of changes in the deterministic components on choice probabilities via the scale parameters. However, this model has received the least attention not only in marketing but also in the economic literature. Further, alternative but not necessarily conflicting interpretations have been given. Allenby and Ginter (1995) used it in their research on consideration sets. Inspired by the work of Hausman and Ruud (1987), this study assumes that for brands under consideration by the consumer, systematic factors have a greater role in determining utilities and therefore, the latter would exhibit less stochastic variation. In transportation research, Bhat (1995) proposed the HEV to account for different variances of unobserved

variables across alternative travel modes. For example, the utility of the train may have more variation than the utility of the car since factors like comfort are less predictable by the commuter. In this manner, uncertainty comes not only from the analyst's observational imperfections but also from the decision maker's imperfect information. Baltas and Doyle (1998) used the HEV to describe brand choice in a frequently bought category. In this study, the model describes unobserved inter and intraindividual fluctuation in preferences that may be unequal across brands. For example, preferences over consumers and suitability over consumption occasions may not vary equally for all competing brands. The origin of the HEV is attributed to Daganzo (1979) who developed an analogous model.

The MNP model (Fig. 1) introduced by Bock and Jones (1968) assumes a multivariate normal distribution and allows for the most general configuration of substitution across products via arbitrary covariance matrices (Hausman and Wise, 1978; Daganzo, 1979; Johnson and Hensher, 1982). The random utilities are allowed to be both correlated and unequal across alternatives, i.e. non-IID. Note that if we set all covariance terms to zero, we get a model similar to HEV with independent random utilities. Likewise, if we further force all variance terms to be equal, we get a model similar to MNL with both independent and identical random utilities. Similarly, if we restrict the covariance terms to be equal within subsets of alternatives but zero between subsets, we get a structure similar to the NMNL model. Unfortunately, despite intellectual appeal, the flexibility of the MNP model comes at the cost of estimation problems. More precisely, there is no closed-form expression for the choice probabilities, which, in principle, require the calculation of an  $(M-1)$ -dimensional integral. Thus, the primary difficulty in using the MNP has been the lack of practical, accurate methods for approximating the choice probabilities when the number of alternatives is large (McFadden, 1981). However, recent advances in estimation technology and, in particular, in simulation-assisted inference (e.g. McFadden, 1989; Hajivassiliou et al., 1996) deal with the tractability problem of the MNP. In practice, the applied researcher still encounters problems such as parameter estimability and computational resources in large problems. Future developments may resolve these issues. Marketing applications have employed simulated maximum likelihood methods (Kamakura and Srivastava, 1984, 1986; Papatla and Krishnamurthi, 1992), McFadden's (1989) method of simulated moments (Chintagunta, 1992a,b; Chintagunta and Honore, 1996; Keane 1997a) and Bayesian approaches (McCulloch and Rossi, 1994). A discussion of the MNP model in connection with recent developments particularly in simulation-based inference is given in Weeks (1997).

#### 4.2. Taste variation

In the preceding section, we have been concerned with models that are able to describe several patterns of non-

IIA unobserved product heterogeneity. This section discusses models that deal explicitly with two related forms of taste variation—interpersonal heterogeneity and intra-personal dynamics.

We begin by considering the treatment of heterogeneity in RU models. From the modeler's standpoint, it makes sense to distinguish between two forms of heterogeneity: heterogeneous tastes for observed attributes and heterogeneous intrinsic tastes for choice alternatives.

More formally, let the systematic utility  $V$  of alternative  $j$  for consumer  $i$  be a function of structural covariates  $x_{ij}$  and an alternative-specific dummy whose coefficient is an alternative-specific intercept  $\alpha_j$ , which can be interpreted as intrinsic brand preference,

$$V_{ij} = \alpha_j + x_{ij}\beta + \varepsilon_{ij}. \quad (6)$$

As alluded to earlier, people may have different tastes for observed characteristics of the alternatives and different sensitivity to marketing variables. Similarly, consumers may have heterogeneous preferences for unmeasured properties of the alternatives which are approximated by nominal variables such as alternative-specific dummies (McFadden, 1980).

This implies that the structural parameters of the model  $\beta$  and also the alternative-specific intercepts  $\alpha_j$  vary over people. Particularly, in working with observations on repeated choices of the sampled individuals (e.g. panel data), this interindividual variation in tastes should not be treated as a random event. Without adjusting for interindividual differences, the model may yield biased estimates of aggregate market response (Chamberlain, 1980; Hsiao, 1986). As depicted in Fig. 1, the literature reflects various approaches to account for taste heterogeneity.

Starting from the most simple strategy, characteristics of the individuals can enter the utility function in cases where a systematic association between individual descriptors and taste variation is assumed. For example, income may be correlated with the importance attached to price. Despite its simplicity, this approach can yield immediate insights into segmentation and targeting as it associates specific individual characteristics with product and attribute preferences. Specific hypotheses can also be tested. For example, a related explanation for the different tastes of consumers with different buying power was introduced by Muellbauer (1975), discussed by Deaton and Muellbauer (1980, p. 262) and applied in purchase data by Allenby and Rossi (1991). Chakraborty et al. (1992) proposed a screening method to identify interactions between demographic variables and attribute coefficients.

Other studies use purchase history as an explanatory variable, given as a function of observed past behavior (e.g. Guadagni and Little, 1983; Krishnamurthi and Raj, 1988) to account for cross-sectional heterogeneity (see also Fader and Lattin, 1993). It is straightforward to employ these two approaches in a standard MNL model. The best way to characterize them is as corrections for the homogeneous model, not as explicit treatments of heterogeneity.

In another related work, consumers are assigned to segments by a deterministic (Currim, 1981; Gensch, 1985) or probabilistic method (Kamakura and Russell, 1989; Bucklin and Gupta, 1992; Kamakura and Russell, 1993; Gupta and Chintagunta, 1994; Kamakura et al., 1996; Mela et al., 1997). Here, the assumption is that taste parameters are constant within segments, which contain homogeneous consumers. The latent-class (LC) model (Kamakura and Russell 1989) assumes that each consumer has a probability of belonging to several latent classes or segments. Thus, the underlying distribution of tastes is assumed to be discrete.

Alternatively, models allow parameters to vary across all households either by estimating parameters for each household, i.e. fixed effects (FE) model or by assuming that these parameters are distributed according to a probability distribution and estimate the parameters of this distribution, i.e. RE model.

There are three main approaches to the FE model. The first available approach involves, at least in theory, estimation of household-specific terms, for instance a constant for each household for each brand. Unfortunately, it quickly becomes impractical as it requires sufficiently long purchase strings and estimation of  $N(M-1)$  alternative-specific intercepts where  $N$ , number of households and  $M$ , number of choice alternatives.

An alternative treatment of heterogeneity in intrinsic alternative preferences has been developed by Chamberlain (1980) and demonstrated in marketing data by Jones and Landwehr (1988). Chamberlain proposed a conditional maximum likelihood method that yields consistent estimates of  $\beta$  by conditioning on sufficient statistics of the intrinsic alternative preferences  $\alpha_{ij}$ . Note that the conditional method takes out of the likelihood function the parameters  $\alpha_{ij}$  and does not estimate them. Its usefulness lies in the consistent estimation of  $\beta$ . Household-level probabilities are not inferred (Chintagunta et al., 1991). As the number of alternatives increases, the conditional ML becomes quite impractical.

A third approach to the FE specification is Bayesian (Rossi and Allenby, 1993). In this technique, the individual purchase string updates pooled parameter estimates to form individual-level counterparts. The prior information ensures, in essence, that all probabilities are positive and resolves the problem of unidentified parameters when some alternatives are never bought by the household.

The RE approach can take two primary forms. The first is a parametric form in which tastes are distributed according to some predetermined continuous distributions such as normal, lognormal, or gamma (Steckel and Vanhonacker, 1988; Gonul and Srinivasan, 1993; Allenby and Lenk, 1995; Kim et al., 1995). Using a continuous distribution, we allow tastes (parameters) to vary across the population according to the specific distribution function and estimate its mean and variance. Attention should be given to the empirical implications of the distributional assumptions. An incorrect probability distribution would result in biased

estimates (Heckman and Singer, 1984). Similarly, the distribution that is invoked to characterize taste heterogeneity should agree with theory about consumer tastes. For example, Gonul and Srinivasan (1993), McCulloch and Rossi (1994), and Allenby and Lenk (1995) use a normal distribution for the price parameter which forces some households to have positive price coefficient. Since economic theory dictates a negative price coefficient, an appropriate distribution with restricted range should be employed (Kim et al., 1995; Brownstone and Train 1996; Train, 1998). For example, the lognormal distribution for the coefficient of the negative of price can do this job.

The second approach to RE modeling is semiparametric. It approximates the underlying distribution of tastes with a discrete probability distribution. The location and probability masses associated with the support points are estimated empirically (Chintagunta et al., 1991; Chintagunta, 1992a; Jain et al., 1994; Chintagunta and Honore, 1996).

Note that the main difference between the semiparametric-RE model and the LC model is one of interpretation. Each support point of the former can be regarded as an LC and similarly each LC can be interpreted as a support point. The difference is that the LC model postulates a finite set of segments of homogeneous members while the RE model, with a discrete distribution, postulates a continuous distribution of tastes that is approximated by a finite set of support points.

A strong pattern in the above studies is that the average impact of policy variables on choice is greater than the level identified by models ignoring heterogeneity (see, e.g. Chintagunta et al., 1991; Chintagunta, 1992a; Gonul and Srinivasan, 1993; Chintagunta and Honore, 1996). Loosely speaking, controlling for interindividual variation increases the explanatory power of non-stochastic factors included in the model. Dealing with heterogeneity may also afford insights for segmentation and fine-tuning of marketing efforts.

Having dealt with issues of interpersonal heterogeneity, we now turn to intrapersonal dynamics. As mentioned above, we shall consider state dependence and inertia as two forms of dynamic choice.

An important empirical problem is the confounding of heterogeneity and state dependence. Time-invariant taste differences over people (i.e. heterogeneity) and effects of the past experiences on current utilities (i.e. state dependence) lead to choice behavior that exhibits persistence over time. Loosely speaking, current behavior conveys information for past choices of the same individual because of two underlying factors: common tastes and dynamic links. Without controlling for heterogeneity, past choices may appear to be a determinant of current behavior solely because they are a proxy for temporally persistent heterogeneous preferences (e.g. Heckman, 1981a,b; Hsiao, 1986). Similarly, if state dependence does exist and is neglected, the degree of time-invariant heterogeneity will be overstated (Keane, 1997a).

From a managerial standpoint, distinguishing heterogeneity from state dependence is important because the degree of the latter determines the long-term results of short-term policies. State dependence implies that short-term activities have long-term potential since there exists a clear causal link between current and future choice behavior (see, e.g. Roy et al., 1996).

The explicit treatment of unobserved heterogeneity makes FE and RE models (e.g. Chintagunta et al., 1991) particularly well-suited for determining dynamic choice in panel data. In practice, state dependence is often accommodated in RU models by making the systematic utility component a function of past choices. This is usually done by introducing a dummy for lagged purchase (Jones and Landwehr, 1988) or a variable summarizing the entire choice history (see Keane, 1997a). Another possibility is to let attributes of previously selected alternatives to influence current attribute tastes (Erdem, 1996).

In addition, current utilities may depend not only on past choices but also on their past values (Heckman, 1981a; Roy et al., 1996). Economists have called this phenomenon, habit persistence. Stated intuitively, the current utility of a brand depends not only on prior buying behavior but also on prior propensities to buy. This short-run inertia in preferences may lead to serially correlated disturbances even in the presence of lagged choice variables which account for true state dependence. Therefore, the empirical detection of habit-persistence according to Heckman (1981a) requires control for state dependence.

As alluded to earlier, serial correlation can arise out of unobserved heterogeneity, state dependence or short-run inertia in preferences (Roy et al., 1996; Keane, 1997a,b). A first attempt at the joint determination of heterogeneity, state dependence, and habit has been made by Roy et al. (1996). This study also provides an instructive discussion of the issues involved in dynamic models. Keane (1997a) develops another quite comprehensive framework that admits structural state dependence and autoregressive error terms while controlling for very rich heterogeneity structures.

#### 4.3. Heterogeneous choice sets

This research stream suggests that consumers may not consider all the available alternatives before making a choice. Thus, choice set may be idiosyncratic for each consumer reflecting a multistage choice process. The economic explanation for this restriction on alternative evaluation is that consumers continue to search for information as long as the returns from that search exceed the respective costs. From a psychological viewpoint, this reduction is viewed as a way to cope with complexity by filtering the alternatives using simple criteria before the detailed evaluation of the reduced set.

The more specific question as to whether choice set effects should be incorporated in RU models does not have a universally accepted answer. For example, Horowitz and

Louviere (1995) argue that choice sets reflect preferences, which are normally described by the utility function.

The relevant literature reflects various approaches to choice set effects. Gensch (1987) hypothesizes an attribute processing stage that determines the options forming the relevant consideration set. Hauser and Wernerfelt (1990) proposed a model where set formation is a trade-off between costs and benefits for including and excluding alternatives. Silk and Urban (1978), Roberts and Lattin (1991), and Horowitz and Louviere (1995) specified sets on the basis of direct consumer reports. Siddarth et al. (1995) define the choice set as the set of options which has been chosen at least once during a time interval. Bronnenberg and Vanhonacker (1996) develop a model that incorporates the probability of each alternative belonging to a consumer's choice set. Ben-Akiva and Boccara (1995) uses information both from consumer reports and observed choice to infer probabilistic choice set generation. Allenby and Ginter (1995) include choice set effects in a single-stage model and examine the influence of merchandising variables. Andrews and Srinivasan (1995) enumerate all possible choice sets and then model the probability that each of them is the choice set for a particular consumer. The current trend is towards probabilistic formulations which view choice sets as latent, unobservable constructs that cannot be imputed with certainty on the basis of observational data. Shocker et al. (1991) and Ben-Akiva and Boccara (1995) note that this implies a more realistic approximation of individual choice behavior.

## 5. RU models and experimental data

RU models have been applied primarily to observational data such as scanner and survey data. This is the typical application for any econometric model, with observations from natural "experiments" provided by the market. Despite the appeal of real-world data, analysts might want to collect experimental choice data for specific projects where observational data are less revealing, for example, new product design (McFadden, 1986; Louviere, 1992).

In experimental choice analysis (ECA), respondents are asked to choose among attribute bundles (e.g. Louviere and Hensher, 1983; Louviere and Woodworth, 1983; Hensher, 1984; Louviere, 1988; Louviere and Batsell, 1991; Chakraborty et al., 1992; Elrod et al., 1992). Then, an RU model is invoked to estimate the mapping from the experimental conditions (attributes of the alternatives and decision makers) to the results (choices). In this respect, the nature of ECA approximates real market behavior better than traditional conjoint methods where there is no corresponding market behavior to the usual rating and ranking tasks (see McFadden, 1986; Louviere, 1994).

Notice that in ECA, all attributes characterizing the hypothetical multiattribute alternatives are known a priori

to the researcher. Therefore, there is less scope for unobserved attributes to influence utilities. Loosely speaking, the stochastic component here deals more with unobserved heterogeneity in decision makers and less with unobserved heterogeneity in alternatives. Unexplained variation at the individual level can be viewed as noise that arises out of confusion, indifference, and inattention and therefore can be treated as a pure chance event that is randomly distributed (Hensher, 1984). The emphasis on taste heterogeneity can also be attributed to the character of the typical experimental task where each respondent is offered a series of choices. Hence, the RU model should account for correlation in unobserved utility over repeated choices by each heterogeneous individual (Revelt and Train, 1997). This argument however should not be over-interpreted. Certain design factors (e.g. brand name) may convey information about various objectives or perceived characteristics and induce unobserved heterogeneity in alternatives.

Finally, a related approach introduced by Beggs et al. (1981) and Chapman and Staelin (1982) is to use an ordered logit model on ranked individual data which allows, in principle, more information to be gathered from respondents but involves other problems (Ben-Akiva et al., 1992). For forceful statements of ECA along with a detailed discussion of several practical issues, the interested reader is referred to Louviere (1994).

## 6. Some practical issues

The literature has given little attention to some practical, yet important issues encountered in applied work. One such problem is the definition of the choice alternative. In most marketing applications, the elemental alternative has been the brand. However, a brand normally covers several variants, e.g. different flavors, formulas, etc. Assuming away most of these details may yield poor policy implications. Aggregation of alternatives may also introduce an aggregation bias that increases with the heterogeneity of elemental alternatives (Parsons and Kealy, 1992). Recently, Fader and Hardie (1996) raised this issue and proposed a way of using detailed attribute information in RU models. The issue of alternative aggregation is directly related to the definition of the alternative set. In a subcategory, for example decaffeinated instant coffee, there is less attribute variation than in the coffee product class. In some cases, the only measurable attributes that vary across alternatives may be merchandising variables and brand name.

When the number of alternatives gets too large, the alternative set may not be manageable. A solution introduced by McFadden (1978) is to randomly sample from the full alternative set and treat the consumer choice as having come from the reduced set. This procedure exploits the IIA property and leads to consistent inference with the MNL



(Parsons and Kealy, 1992) but it is not certain how it works with other models. Another strategy to reduce the computational burden of a large number of alternatives is to sample from the full alternative set with unequal selection probabilities which reflect the importance of alternatives (Ben-Akiva and Lerman, 1985, pp. 261–269). Alternatively, the researcher can cluster items sharing the same attribute and estimate a NMNL model (Kannan and Wright, 1991) or assume heterogeneous choice sets and employ one of the procedures discussed in the previous section.

A related empirical issue is how to deal with a large number of explanatory variables in discrete choice models—a problem often encountered in working with individual descriptors. As alluded to earlier, Chakraborty et al. proposed a method for empirically selecting significant variables in choice models.

Another important empirical issue noted by Krishnamurthi and Raj (1991) and analyzed empirically by Kim and Rossi (1994) relates to the exclusion of infrequent purchasers because of estimation or other requirements. Such sample selection rules may induce a heavy-user bias (Kim and Rossi, 1994). Most research on scanner data also tends to drop small brands for lack of observations, or to simplify the analysis. While such sample selection practices could be justified for illustrating a new model or approach, they are less acceptable in applied, business settings where a complete picture of the market is required.

Finally, a general concern relates to overall model practicality. As our discussion illustrates, recent developments have increased model complexity and made estimation, interpretation, and forecasting less straightforward. Some specifications are still rather impractical. The issue can be viewed as the common dilemma between simplicity and flexibility. There is no universal answer to this question as it depends on one's rate of exchange between the two criteria.

## 7. Concluding remarks

In this paper, we have been concerned with RU models for discrete choice. The theoretical framework laid out in the outset of our discussion allowed several issues and models to be presented and interpreted. It is evident that RU models have proved useful methods for assessing the effects of marketing mix, guiding new product development, and forecasting. Part of their current popularity stems from recent advances in the collection of behavioral data through retail checkouts. Besides packaged goods, RU models can also be used to analyze discrete choice behavior in services, distribution, education, and other business settings involving optimizing decisions. Finally, the models discussed in this paper can be incorporated in more general demand systems that attempt the independent (e.g. Gupta, 1988) or joint (e.g. Chiang, 1991; Chintagunta, 1993; Baltas 1998) determination of discrete and continuous decisions such as which brand and how many units to buy.

## Acknowledgments

Because of space limitations some contributions have not been given the attention that they deserve. The authors thank two anonymous JBR reviewers and the editors for helpful comments.

## References

- Allenby GM, Ginter JL. The effects of in-store displays and feature advertising on consideration sets. *Int J Res Mark* 1995;12:67–79.
- Allenby GM, Lenk PJ. Reassessing brand loyalty, price sensitivity, and merchandising effects on consumer brand choice. *J Bus Econ Stat* 1995;13:281–9.
- Allenby GM, Rossi P. Quality perceptions and asymmetric switching between brands. *Mark Sci* (Summer) 1991;10:185–204.
- Anderson SP, de Palma A, Thisse J-F. *Discrete Choice Theory of Product Differentiation*. Cambridge, MA: MIT Press, 1992.
- Andrews RL, Srinivasan TC. Studying consideration effects in empirical choice models using scanner panel data. *J Mark Res* (February) 1995;32:30–41.
- Baltas G. An integrated model of category demand and brand choice. *J Mark Res Soc* 1998;40:295–306.
- Baltas G, Doyle P. A flexible model for consumer choice in packaged goods categories. *J Mark Res Soc* 1998;40:141–53.
- Baltas G, Doyle P, Dyson P. A model of consumer choice for national vs. store brands. *J Oper Res Soc* 1997;48:988–95.
- Beggs S, Cardell S, Hausman J. Assessing the potential demand for electric cars. *J Econometrics* 1981;16:1–19.
- Ben-Akiva M. *Structure of Passenger Travel Demand Models*. PhD dissertation. Department of Civil Engineering, MIT, 1973.
- Ben-Akiva M, Boccara B. Discrete choice models with latent choice sets. *Int J Res Mark* 1995;12:9–24.
- Ben-Akiva M, Lerman S. *Discrete Choice Analysis: Theory and Application to Predict Travel Demand*. Cambridge, MA: MIT Press, 1985.
- Ben-Akiva M, Morikawa T, Shiroishi F. Analysis of the reliability of preference ranking data. *J Bus Res* 1992;24:149–64.
- Bhat CR. A heteroscedastic extreme value model of intercity travel mode choice. *Transp Res B* 1995;29:471–83.
- Bock RD, Jones LV. *The Measurement and Prediction of Judgement and Choice*. San Francisco, CA: Holden-Day, 1968.
- Bronnenberg BJ, Vanhonerack WR. Limited choice sets, local price response, and implied measures of price competition. *J Mark Res* (May) 1996;33:163–73.
- Bronnenberg BJ, Wathieu L. Asymmetric promotion effects and brand positioning. *Mark Sci* (Fall) 1996;15:379–94.
- Brownstone D, Train K. Forecasting New Product Penetration with Flexible Substitution Patterns. Working Paper, Department of Economics, University of California, Berkeley, 1996.
- Buckley PG. Nested multinomial logit analysis for a hierarchical choice model. *J Bus Res* (September) 1988;17:133–54.
- Bucklin RE, Gupta S. Brand choice, purchase incidence and segmentation: an integrated modeling approach. *J Mark Res* 1992;29:201–15.
- Chakraborty G, Woodworth G, Gaeth GJ, Ettenson R. Screening for interactions between design factors and demographics in choice-based conjoint. *J Bus Res* (March) 1992;24:115–33.
- Chamberlain G. Analysis of covariance with qualitative data. *Rev Econ Stud* 1980;47:225–38.
- Chapman R, Staelin R. Exploiting rank ordered choice set data within the stochastic utility model. *J Mark Res* 1982;18:288–301.
- Chiang J. A simultaneous approach to whether, what and how much to buy questions. *Mark Sci* 1991;10:297–315.
- Chintagunta PK. Heterogeneity in nested logit models: an estimation approach and empirical results. *Int J Res Mark* 1992;9:161–75.

- Chintagunta PK. Estimating a multinomial probit model of brand choice using the method of simulated moments. *Mark Sci* (Fall) 1992;11:386–407.
- Chintagunta PK. Investigating purchase incidence, brand choice and purchase quantity decisions of households. *Mark Sci* (Spring) 1993;12:184–208.
- Chintagunta PK, Honore BE. Investigating the effects of marketing variables and unobserved heterogeneity in a multinomial probit model. *Int J Res Mark* 1996;13:1–15.
- Chintagunta PK, Jain DC, Vilcassim NJ. Investigating heterogeneity in brand preferences in logit models for panel data. *J Mark Res* (November) 1991;28:417–28.
- Currim IS. Using segmentation approaches for better prediction and understanding from consumer choice models. *J Mark Res* (August) 1981;18:301–9.
- Daganzo C. *Multinomial Probit: The Theory and Applications to Demand Forecasting*. New York: Academic Press, 1979.
- Deaton A, Muellbauer J. *Economics and Consumer Behavior*. Cambridge: Cambridge Univ. Press, 1980.
- Deighton J, Henderson CM, Nelsin SA. The effects of advertising on brand switching and repeat purchasing. *J Mark Res* (February) 1994;34:28–43.
- Dubin JA. A nested logit model of space and water heat system choice. *Mark Sci* 1986;5:112–24.
- Elrod TS, Louviere J, Kavey KS. An empirical comparison of ratings-based and choice-based conjoint models. *J Mark Res* 1992;29:368–77.
- Erdem T. A dynamic analysis of market structure based on panel data. *Mark Sci* 1996;15:359–78.
- Fader PS, Hardie BGS. Modeling consumer choice among SKUs. *J Mark Res* 1996;33:442–52.
- Fader PS, Lattin JM. Accounting for heterogeneity and nonstationarity in a cross-sectional model of consumer purchase behavior. *Mark Sci* (Summer) 1993;12:304–17.
- Gensch DH. Empirically testing a disaggregate choice model for segments. *J Mark Res* (November) 1985;22:462–7.
- Gensch DH. A two-stage disaggregate attribute choice model. *Mark Sci* (Summer) 1987;6:223–39.
- Gensch DH, Recker WW. The multinomial, multiattribute logit choice model. *J Mark Res* (February) 1979;16:124–32.
- Gonul F, Srinivasan K. Modeling multiple sources of heterogeneity in multinomial logit models: methodological and managerial issues. *Mark Sci* (Summer) 1993;12:213–29.
- Guadagni PM, Little JD. A logit model of brand choice calibrated on scanner data. *Mark Sci* 1983;2:203–38.
- Gupta S. Impact of sales promotions on when, what, and how much to buy. *J Mark Res* (November) 1988;25:342–55.
- Gupta S, Chintagunta PK. On using demographic variables to determine segment membership in logit mixture models. *J Mark Res* (February) 1994;31:128–36.
- Hajivassiliou V, McFadden D, Ruud P. Simulation of multivariate normal rectangle probabilities and their derivatives. *J Econometrics* 1996;72:85–134.
- Hardie BGS, Johnson EJ, Fader PS. Modeling loss aversion and reference dependence effects on brand choice. *Mark Sci* (Fall) 1993;12:378–94.
- Hauser JR, Urban GL. A normative methodology for modeling consumer response to innovation. *Oper Res* 1977;25:579–619.
- Hauser JR, Wernerfelt B. An evaluation cost model of consideration sets. *J Consum Res* (March) 1990;16:393–408.
- Hausman J, Ruud P. Specifying and testing econometric models for rank-order data. *J Econometrics* 1987;34:83–104.
- Hausman J, Wise DA. A conditional probit model for qualitative choice: discrete decisions recognising interdependence and heterogeneous preferences. *Econometrica* 1978;46:403–26.
- Heckman JJ. Statistical models for discrete panel data. In: Manski CF, McFadden D, editors. *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, MA: MIT Press, 1981. pp. 114–78.
- Heckman JJ. The incidental parameters problem and the problem of initial conditions in estimating a discrete time–discrete data stochastic process. In: Manski CF, McFadden D, editors. *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, MA: MIT Press, 1981. pp. 179–95.
- Heckman JJ, Singer BA. Method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica* 1984;52:271–320.
- Hensher DA. Achieving representativeness of the observable component of the indirect utility function in logit choice models: an empirical revelation. *J Bus* 1984;57:265–80.
- Horowitz JL, Louviere JJ. What is the role of consideration sets in choice modeling. *Int J Res Mark* 1995;12:39–54.
- Hsiao C. *Analysis of Panel Data*. Cambridge: Cambridge Univ. Press, 1986.
- Jain CJ, Vilcassim NF, Chintagunta PK. A random-coefficients logit brand-choice model applied to panel data. *J Bus Econ Stat* 1994;12:317–28.
- Johnson EJ, Hensher D. Application of multinomial probit to a two-period panel data set. *Transp Res A* 1982;16:457–64.
- Jones MJ, Landwehr JT. Removing heterogeneity bias from logit model estimation. *Mark Sci* (Winter) 1988;7:41–59.
- Kamakura WA, Kim B-D, Lee J. Modeling preference and structural heterogeneity in consumer choice. *Mark Sci* 1996;15:152–72.
- Kamakura WA, Russell GJ. A probabilistic choice model for market segmentation and elasticity structure. *J Mark Res* (November) 1989;26:379–90.
- Kamakura WA, Russell GJ. Measuring brand value with scanner data. *Int J Res Mark* 1993;10:9–22.
- Kamakura WA, Srivastava RK. Predicting choice shares under conditions of brand interdependence. *J Mark Res* (November) 1984;21:420–34.
- Kamakura WA, Srivastava RK. An ideal-point probabilistic choice model for heterogeneous preferences. *Mark Sci* (Summer) 1986;5:199–218.
- Kannan PK, Wright GP. Modeling and testing structured markets: a nested logit approach. *Mark Sci* 1991;10:58–82.
- Keane MP. Modeling heterogeneity and state dependence in consumer choice behavior. *J Bus Econ Stat* 1997;15:310–11.
- Keane MP. Current issues in discrete choice modeling. *Mark Lett* 1997;8:307–22.
- Kim B-D, Blattberg RC, Rossi PE. Modeling the distribution of price sensitivity and implications for optimal retail pricing. *J Bus Econ Stat* 1995;13:291–303.
- Kim B-D, Rossi PE. Purchase frequency, sample selection and price sensitivity: the heavy user bias. *Mark Lett* 1994;5:57–68.
- Krishnamurthi L, Raj SP. A model of brand choice and purchase quantity price sensitivities. *Mark Sci* (Winter) 1988;7:1–20.
- Krishnamurthi L, Raj SP. An empirical analysis of the relationship between brand loyalty and consumer price elasticity. *Mark Sci* 1991;10:172–83.
- Lattin J, Bucklin RE. Reference effects of price and promotion on brand choice behavior. *J Mark Res* (August) 1989;26:299–310.
- Lattin JM. A model of balanced choice behavior. *Mark Sci* 1987;6:48–65.
- Louviere J. Conjoint analysis. In: Bagozzi RP, editor. *Advanced Methods of Marketing Research*. Cambridge, MA: Blackwell, 1994. pp. 223–59.
- Louviere J, Hensher DA. Using discrete choice models with experimental design data to forecast consumer demand for a unique cultural event. *J Consum Res* (December) 1983;10:348–61.
- Louviere J, Woodworth G. Design and analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data. *J Mark Res* 1983;20:350–67.
- Louviere JJ. Conjoint modeling of stated preferences: a review of theory, methods, recent developments and external validity. *J Transp Econ Policy* 1988;10:93–119.
- Louviere JJ. Experimental choice analysis: introduction and overview. *J Bus Res* 1992;24:89–95.
- Louviere JJ, Batsell RR. Experimental analysis of choice. *Mark Lett* 1991;2:199–214.
- Manski CF. The structure of random utility models. *Theory Decis* 1977;8:229–54.
- McCulloch R, Rossi PE. An exact likelihood analysis of the multinomial probit model. *J Econometrics* 1994;64:207–40.

- McFadden D. Conditional logit analysis of qualitative choice behavior. In: Zarembka P, editor. *Frontiers in Econometrics*. New York: Academic Press, 1973. pp. 105–42.
- McFadden D. Modelling the choice of residential location. In: Karquist A, et al., editors. *Spatial Interaction Theory and Planning Models*. Amsterdam: North-Holland, 1978. pp. 75–96.
- McFadden D. Modeling the choice of residential location. *Transp Res Rec* 1979;673:72–8.
- McFadden D. Econometric models for probabilistic choice among products. *J Bus* 1980;53:513–29.
- McFadden D. Econometric models of probabilistic choice. In: Manski CF, McFadden D, editors. *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, MA: MIT Press, 1981. pp. 198–272.
- McFadden D. The choice theory approach to market research. *Mark Sci* 1986;5:275–97.
- McFadden D. A method of simulated moments for the estimation of discrete response models without numerical integration. *Econometrica* 1989;57:995–1026.
- Mela CF, Gupta S, Lehmann DR. The long-term impact of promotion and advertising on consumer brand choice. *J Mark Res* (May) 1997;34:248–61.
- Muellbauer J. The cost of living and taste and quality changes. *J Econ Theory* 1975;10:269–83.
- Papatla P, Krishnamurthi L. A probit model of choice dynamics. *Mark Sci* (Spring) 1992;11:189–206.
- Parsons GR, Kealy MJ. Randomly drawn opportunity sets in a random utility model of lake recreation. *Land Econ* 1992;68:93–106.
- Punj GN, Staelin R. The choice process for graduate business schools. *J Mark Res* 1978;15:588–98.
- Revelt D, Train K. Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level, working paper, Department of Economics, University of California, Berkeley, 1997.
- Roberts JH, Lattin JM. Development and testing of a model of consideration set composition. *J Mark Res* (November) 1991;28:429–40.
- Rossi PE, Allenby GM. A Bayesian approach to estimating household parameters. *J Mark Res* (May) 1993;30:171–82.
- Roy R, Chintagunta PK, Haldar S. A framework for investigating habits, “The Hand of the Past,” and heterogeneity in dynamic brand choice. *Mark Sci* 1996;15:280–99.
- Shocker AD, Ben-Akiva M, Boccara B, Negungedi P. Consideration set influences on consumer decision making and choice: issues, models, and suggestions. *Mark Lett* 1991;2:181–97.
- Siddarth S, Bucklin RE, Morrison DG. Making the cut: analyzing choice set restriction in scanner panel data. *J Mark Res* (August) 1995;32:255–66.
- Silk AJ, Urban GL. Pre-test-market evaluation of new packaged goods: a model and measurement methodology. *J Mark Res* 1978;15:171–91.
- Steckel JH, Vanhonnacker WR. A heterogeneous conditional logit model of choice. *J Bus Econ Stat* 1988;6:391–8.
- Swait J, Erdem T, Louviere J, Dubelaar C. The equalization price: a measure of consumer-perceived brand equity. *Int J Res Mark* 1993;10:23–45.
- Tellis GJ. Advertising exposure, loyalty and brand purchase: a two-stage model of choice. *J Mark Res* 1988;25:134–44.
- Train K. *Qualitative Choice Analysis*. Cambridge, MA: MIT Press, 1986.
- Train KE. Recreation demand models with taste differences over people. *Land Econ* 1998;74:230–9.
- Weeks M. The multinomial probit model revisited: a discussion of parameter estimability, identification and specification testing. *J Econ Surv* 1997;11:297–320.
- Winer RS. A reference price model of brand choice for frequently purchased products. *J Consum Res* (September) 1986;13:250–6.