A New Brand Choice Model Incorporating
A Choice Set Formation Process

Wilfried R. Vanhonacker

Series No. MKTG 94.001
Revised, December 1993
Revised, October 1994
A New Brand Choice Model Incorporating
A Choice Set Formation Process

Wilfried R. Vanhonacker *

Revised, December 1993
Revised, October 1994

* Professor of Marketing, INSEAD, Fontainebleau, France and the Hong Kong University of Science & Technology, Hong Kong. This research is part of a single-source model development project for A.C. Nielsen and Euromedia (France). The paper has benefited from the comments of Niraj Dawar, Lydia Price, Bart Bronnenberg, and three anonymous reviewers.
A New Brand Choice Model Incorporating A Choice Set Formation Process

ABSTRACT

This paper develops an individual-level choice model which embeds a rational choice set generation process. Choice set generation is conceptualised as a process in which the consumer makes uncertain judgments about the ability of brands to meet the consumption or usage goal motivating the choice decision. Brand selection is modeled as the joint maximization of that ability and the brand's relative value (or relative attractiveness) which the individual consumer assesses at the time of choice and independently of any usage or consumption goal. Derivations are provided for a flexible and parsimonious representation of the choice probabilities which are not characterized by the IIA property. Empirical results using scanner data and survey-based awareness data for a single product category provide support for the model. Apart from enriching the behavioral insight into the brand choice decision, the results show that not recognising unequal substitutability among brands arising from their differential ability to meet the usage/consumption goal leads to non-optimal marketing mix decisions.

Key Words: Discrete Choice Models; Choice Set Generation; Typicality
INTRODUCTION

Parsimonious models of the individual consumer's choice process have an extensive tradition in marketing. Describing choice as a rational process has helped in our understanding and explanation of observed choice outcomes. The utility maximizing choice rule has played a central role in that regard. The rational process of utility maximization leads the individual consumer to choose the alternative with the highest utility. Hence, choice results from a relative comparison conditional on a choice set which consists of all available alternatives actually considered. Some authors have referred to this set as the consideration set after Wright and Barbour (1977). Choice set generation then refers to the process of defining the "considered subset" (Simon 1955).

Until recently, little attention has been devoted in the marketing brand choice literature to the process by which such choice sets are formed. Indeed, most operationalizations define choice as the selection of a particular brand from all those available in the market place (see, e.g. Gensch and Recker 1979, Guadagni and Little 1983, Gupta 1988, Lattin and Bucklin 1989, Chintugunta, Jain and Vilcassin 1991, Fader, Lattin and Little 1993, Fader and Lattin 1993). This tradition of equating choice set with the set of all available alternatives is rather surprising given conflicting behavioral evidence which indicates that consumers often select from a limited number of items (see, e.g., Wright and Barbour 1977, Bettman 1979, 1987, Payne 1982). This disparity between choice models and empirical evidence is troubling in light of extensive empirical and theoretical work in the transportation literature which argues that the implications of choice set misspecification can be quite severe (see, e.g., Stopher 1980, Williams and Ortuzar 1982, Swait 1984, Swait and Ben-Akiva 1986).

Recognizing these dangers, recent work in marketing has followed the transportation literature (see, e.g., Ben-Akiva 1977, Swait 1984, Pitschke 1980, Swait and Ben-Akiva 1987, Boccara 1989, Ben-Akiva and Boccara 1990) in focusing on the issue of choice or consideration set generation (Roberts 1989, Nedungadi 1987, 1990, Shocker et. al. 1991), and initial attempts to integrate the choice set generation process into discrete consumer choice models have appeared (Roberts and Lattin 1991, Hauser and Wernerfelt 1989, 1990). Hauser and Wernerfelt (1989, 1990), for example, formalize the process of choice set formation in a rational fashion using search theory (analogous to Meyer 1982 and Richardson 1982). Building on the information search paradigm, Roberts and Lattin (1991) provide an operationalization, similar to that of Ben-Akiva and Lerman (1985), of the incremental benefit derived from including another alternative in the choice set. This representation of choice set generation, which is calibrated separately across consumers, is integrated into a multinominal logit model of choice.
While improving on earlier models which completely lack a choice set specification, these latter modeling efforts still fail to recognize an important aspect of the choice set generation process. As discussed in Shocker et. al. (1991), the usage or consumption occasion motivating the choice is likely to influence the nature and composition of the choice set. Indeed, Ratneshwar and Shocker (1991) argue that usage contexts act as environmental constraints that help the individual consumer define ends and goals, and, hence, limit the nature of the brands that can help achieve consumption objectives. The notion of consumption objectives, or goals, implies that the consumer will judge the extent to which brands will enable goal achievement. This judgment is central to the choice set generation process.

One approach which recognizes the inherent characteristics of the choice set generation process is to conceptualize it as a goal-derived categorization process (Barsalou 1983, 1985). The goal-derived categorization paradigm, invoking the construct of typicality, has been well accepted in consumer behavior (see, e.g., Sujan 1985, Nedungadi and Hutchinson 1985, Ward and Loken 1986, Loken and Ward 1987, Nedungadi 1990, Nedungadi and Kanetkar 1992). Typicality refers to the phenomenon that not all members of a category are equally representative of that category. In goal-derived categorization, typicality of an object depends on its ability to meet a goal, with ability varying across category members. Choice set generation then can be viewed as a categorization task with the usage or consumption occasion motivating the choice as the goal (Troye 1984). Here, typicality would refer to the ability of the choice set members to help achieve the consumption or usage goal. As there is evidence that consumers organize categories around typical or exemplar brands (Sujan 1985), choice set formation then arises from typicality judgments relative to a “prototypical” or “ideal” brand for the consumption goal.

In this paper, a rational operationalization of the choice set generation process is provided which explicitly recognizes possible uncertainty in the consumer's typicality judgments arising from a lack of familiarity with the brand and/or the usage/consumption occasion motivating the choice decision. Previous attempts to model choice set generation have not recognized this element of uncertainty. As the literature on decisions under uncertainty amply illustrates, it is very important to recognize this uncertainty as it affects the choice rule with the consumer attempting to limit the risk implied. Furthermore, in our specific context here, this uncertainty gives rise to a probabilistic assessment of choice set membership for all brands on the part of the consumer. Specifically, the individual consumer will make an uncertain judgment about whether each brand available is worth considering as a choice alternative. Conceptually, this is drastically different from previous probabilistic definitions of choice sets (see, e.g., Manski 1977, Swait and Ben-Akiva 1987, Ben-Akiva and Boccara 1990, Horowitz and Louviere 1991, Siddarth, Bucklin, and Morrison 1993). In this stream
of literature, the probabilistic definition of choice set arises exclusively from uncertainty on the part of the modeler about the exact content of the discrete but latent choice set used by the individual consumer. Even the uncertainty introduced in Roberts and Lattin (1991) only represents modeler’s uncertainty about the exact utility threshold the consumer uses to identify choice options. Our conceptualisation of the choice set generation process at the individual level is a clear departure from this research.

In operationalising the brand choice decision, the individual consumer is assumed to integrate his/her uncertain brand typicality judgments (which are consumption/usage-context dependent) with brand evaluations which capture the relative attractiveness of the brand offer but are consumption/usage-context independent. Such include, for example, evaluations of temporary price cuts or purchase quantity promotions (e.g., two-for-the-price-of-one) which may make a brand more attractive independent from the usage/consumption occasion motivating the choice. The choice rule the rational consumer will use in integrating both aspects is one of joint maximization. Specifically, the consumer is assumed to jointly maximize choice set membership based on relative typicality and relative value. The derived choice model is parsimonious and is not characterized by the Independence of Irrelevant Alternatives (IIA) property (Currim 1982).

The contribution of the choice model developed here lies in its comprehensive integration of a behaviorally-motivated choice set generation process in a random utility choice framework. It illustrates the importance of considering the usage context or consumption occasion motivating the brand choice decision and its impact on the rational choice rule. Conceptually, the model combines theories of goal-derived categorization in cognitive psychology with theories of random utility. Operationally, the model combines inherent uncertainty in choice set generation with random utility evaluation inherent in the choice task within an overall rational framework. Managerially, the model provides better insights into brand choice decisions and illustrates that normative decisions on marketing mix elements are conditional on the brand’s likelihood of being included in the choice set versus its relative value.

An empirical illustration using supermarket scanner data supplemented with survey-based awareness data for a single product category provides support for the model. Furthermore, the results indicate that traditional discrete choice models such as the multinominal logit model (Mcfadden 1974) which ignore choice set generation and its implicit brand typicality heterogeneity can provide erroneous normative insights. Specifically focusing on promotion frequency, the results indicate the significant degree to which multinominal logit results would lead store brands to under-promote and name brands to over-promote in the promotion-sensitive category.
The paper is structured as follows. First, the conceptual background of the choice model is discussed. Second, the choice rule and the basic model structure are operationalised. Third, the choice set generation process is operationalised. Fourth, the choice model and its characteristics are discussed. Fifth, an empirical example is reported and discussed in detail. Conclusions and directions for future research follow.

BACKGROUND

Fundamental to the conceptual understanding of the choice model developed in this paper are the brand evaluations the individual consumer will perform in arriving at the choice decision, the choice rule used in making that decision, and the impact of the latent nature of that evaluation process and choice rule on the formalisation of the brand choice decision by the modeler. In this section, we first discuss the brand evaluation process. Subsequently, the choice rule for a rational consumer is derived. Finally, we discuss how latent aspects of the consumer's brand evaluation and choice process lead the modeler to formalise the brand choice decision in a particular manner.

The brand evaluations the individual consumer will make consist basically of two components: one which is conditional on the usage or consumption occasion motivating the choice, and another which is unconditional on that occasion. The brand evaluation component which is conditional on the usage or consumption occasion consists of a subjective assessment of the brand's fit with that occasion. The brand evaluation component which is unconditional on the usage or consumption occasion consists of a subjective assessment of the relative attractiveness of a particular brand irrespective of the envisioned usage or consumption. For example, if a consumer is looking for a healthy snack, he/she may value a low-calorie yogurt brand favorably. On the other hand, budgetary considerations may result in that consumer looking favorably upon another brand whose price is temporarily reduced. Where the favorable evaluation of some yogurt brand because of its low calorie content is conditioned by the envisioned consumption occasion of having a healthy snack, the favorable evaluation of the brand on promotion is entirely independent of any usage or consumption occasion.

Although one could argue that both components are commonly accounted for in the utility evaluations underlying traditional discrete choice models, they are separated here for three main reasons. First, both components play a distinctive role in the evaluation process leading up to the choice decision. Furthermore, because of their distinctive role, there is a specific sequence in which they are evaluated and enter the process. Second, both evaluation
components arise from a task which is different in terms of nature and uncertainty. Specifically, we will argue that the usage-context dependent evaluation is essentially a brand-by-brand evaluation relative to an exemplar (or prototypical) alternative characterised by uncertainty on the part of the individual consumer where the usage-context independent evaluation is a certain but pairwise evaluation between brands. Third, and important for operationalising the choice decision, both components are evaluated independently from one another even though they may share some common determinants. We discuss each of these in some detail now.

With respect to their distinctive roles, the brand evaluation component which is conditional on the usage/consumption occasion assesses the extent to which the individual consumer considers the brand as a choice alternative. In the above example, the consumer judges low-calorie yogurt brands to be consistent with a healthy snack and as a result is more likely to consider those as choice alternatives. Shocker et. al. (1991) and others have argued that the usage or consumption occasion motivating a choice decision is an important determinant in choice set formation. Our argument here is that in identifying likely choice options, the individual consumer will assess the extent to which each available brand can contribute to achieving the envisioned usage/consumption goal. In this research, we follow Troye (1984) and rely on the goal-derived categorisation paradigm (Barsalou 1983, 1985) to conceptualise that assessment. It is evident that such conceptualisation is valid only in the context where brand choices are indeed motivated by future usage/consumption occasions, and we will limit our developments to those instances.

Where the usage-context dependent evaluation results in likelihood of consideration, the usage-context independent brand evaluation makes the brand stand out among all the options. It does not influence the brand's typicality assessment on the part of the individual consumer. In the above example, the price promotion run by one brand does not make that brand more of a healthy snack in the eyes of the consumer; the price promotion makes that brand more attractive because of the good value it provides in light of budget constraints on the part of the consumer. Conceptually, we argue that the conditional evaluation component has a direct influence on choice set generation where the unconditional evaluation component has a direct influence on brand selection. Of course, through its impact on choice set generation, the usage-context dependent evaluation component influences brand selection as well but in a more indirect fashion. Note that our distinction between the evaluation components and their role is in line with Nedungadi's (1990) argument that choice set generation and brand selection play a different role. Furthermore, given consumer behavior research which suggests that choice set generation is done prior to choice (see, e.g., Wright and Barbour 1977, Bettman 1979, 1987, Payne 1982), the conceptualisation adopted
here implies that the usage-context dependent evaluation is done prior to the usage-context independent evaluation.

Apart from their distinctive roles, and perhaps more important in terms of understanding individual brand choice decisions, both brand evaluation components arise from distinctive and independent tasks. As discussed above, the conditional brand evaluation consists of the assessment of the brand's typicality given the consumption/usage goal motivating the choice. At the individual consumer level, that assessment implies uncertainty which can arise from limited knowledge or experience with the brand and/or with the consumption/usage occasion. If, in the above example, the consumer upon doctor's advice has only recently become more health conscious in his/her eating habits, he/she might not be all that confident about how healthy a low-calorie yogurt is. Furthermore, given continuous product reformulations, improvements, and introductions even in established product categories, uncertainty arising from a lack of product-related experiences is likely to arise as well. This uncertainty cannot be ignored as the consumer who experiences it will adapt his/her behavior to it as the uncertainty implies some risk that the brand might not satisfy the usage/consumption goal. Apart from choice rule implications, the uncertainty results in a probabilistic definition of choice set membership. From the perspective of the individual consumer, each brand has some likelihood of being in the choice set, with that likelihood being equal to that consumer's assessment of the brand's typicality (i.e., extent to which it helps satisfying the consumption/usage goal). Note that these assessments imply unequal substitutability among the brands which is a departure from the assumption of equal substitutability among choice alternatives belonging to the choice set made in, for example, the multinomial logit choice model (Ben-Akiva and Lerman 1985). As we will discuss and illustrate later, this unequal substitutability arising from the differential typicality across brands given a specific consumption/usage occasion has important normative implications.

In contrast, the unconditional brand evaluation component implies no uncertainty on the part of the individual consumer. Analogous to assumptions underlying traditional random utility theory (Manski 1977), the individual consumer can make an accurate assessment of the attractiveness of a brand relative to others on offer. In the above example, the individual consumer can evaluate with certainty the monetary gain provided by the brand on promotion.

Although this usage-context independent evaluation of brand value may be influenced by determinants underlying the typicality judgments (such as, e.g., brand preferences), it is done very much independent from the usage-context dependent evaluation. In the example, the evaluation of a yogurt brand's healthiness has little to do with evaluating its monetary substitution value. Nevertheless, both will play a role in the brand choice decision. The
question then arises of what choice rule the individual consumer is going to use to select a brand. Both the brand's typicality and the brand's relative value are important to the consumer. Hence, both independent evaluation components will be integrated in a rational choice rule. We postulate in this research that the individual consumer will jointly maximize relative typicality and relative value. As high relative typicality implies low risk that the brand will not meet the consumption/usage goal, the choice rule can be interpreted as one of jointly maximizing the attractiveness of the offer and minimizing the risk of selecting a brand which will not help in satisfying the usage/consumption goal. This rational choice rule and the basic choice model structure recognising the latent character of both brand evaluation components on the part of modeler are formalised next.

**CONSUMER CHOICE MODEL STRUCTURE**

We denote \( \Pi_i(j \in M^\ell) \) as consumer \( i \)'s subjective assessment of brand \( j \)'s likelihood of belonging to choice set \( M \) for usage/consumption occasion \( \ell \). The nature of that assessment and the role of prototypical brands as reference points in it (together with what is prototypical) is operationalised hereafter. We first proceed with formalising the basic model given the subjective outcome of that assessment. Furthermore, we denote \( V_{ij} \) as consumer \( i \)'s deterministic assessment of brand \( j \)'s value. Limiting ourselves to this component, the relative attractiveness of brand \( j \) to consumer \( i \), \( R_{Ai j} \), can be captured in a "strict utility" model sense (Luce 1959) as

\[
R_{Ai j} = \left( \frac{V_{ij}}{\sum_{k \in J} V_{ik}} \right)
\]

where \( J \) denotes the universal set of all brands. If the individual consumer would only make usage-context independent evaluations and base his/her choice decisions on that using a choice rule of maximizing value, the ratio \( R_{Ai j} \) would denote the relative frequency with which we would observe consumer \( i \) selecting brand \( j \) over repeated choices.

In arriving at a choice decision, however, both evaluation components are important to the consumer. On one hand, he/she prefers \( \Pi_i(j \in M^\ell) \) to be as close to one as possible; on the other hand, he/she prefers \( V_{ij} \) (or \( R_{Ai j} \)) to be as large as possible. Given that both evaluation components are probabilities (i.e., \( \Pi_i(j \in M^\ell) \) and \( R_{Ai j} \), respectively) and are independent, we argue that the brand preference order on the part of the individual consumer is determined by the joint probability \( \left[ \Pi_i(j \in M^\ell) \cdot R_{Ai j} \right] \). Accordingly, the rational choice rule individual consumer \( i \) will use is one of maximizing \( \left[ \Pi_i(j \in M^\ell) \cdot R_{Ai j} \right] \) over all brands \( j \in J \). From a modeler's perspective, however, both components are unobservable. In other
words, the modeler is uncertain about the evaluations the individual consumer has made as well as the subjective probability outcomes of those evaluations. In formalising the model, we have to integrate modeler uncertainty in both evaluation components and derive the choice probabilities consistent with those components and the choice rule used by the individual consumer.

As discussed above, \( \Pi_i(j \in M^f) \) arises from typicality judgments which contain an element of uncertainty. From a modeler's perspective, they are latent and, hence, have an additional element of uncertainty. We will operationalise shortly the process model of typicality evaluations and integrate at that point uncertainty which will capture both the individual consumer's uncertainty as well as the modeler's uncertainty. The outcome will be \( P_i(j \in M^f) \) which is the modeler's assessment of the individual consumer's \( \Pi_i(j \in M^f) \).

For integrating modeler uncertainty into the usage-context independent evaluation component, we follow the traditional formalisation used in random utility theory (Manski 1977). Specifically, we define

\[
U_{ij} = V_{ij} + \mu_{ij} \tag{1}
\]

where \( U_{ij} \) denotes the value of brand \( j \) to consumer \( i \). Random component \( \mu_{ij} \) captures modeler uncertainty given that the consumer's deterministic assessment captured in \( V_{ij} \) is unobservable to the modeler. Assuming that the random components are independent and identically distributed (i.i.d.) according to a Type-1 Extreme Value distribution, the relative value of the brands as assessed by the modeler is captured by a logit model, or

\[
P_{i(j)} = \left[ \exp(V_{ij}) / \sum_{k \in J} \exp(V_{ik}) \right] . \tag{2}
\]

This expression captures the modeler's assessment of the consumer's relative attractiveness evaluation, or \( RA_{ij} \).

Applying the choice rule of joint maximization implies a modeled preference order determined by \( P_i(j \in M^f), P_{i(j)} \) which provides the choice probabilities

\[
P_{ij}^f = \left[ P_i(j \in M^f), P_{i(j)} \right] / \left[ \sum_{k \in J} P_i(k \in M^f), P_{i(k)} \right]
\]

10
which, given expression (2), equal

\[
P^\ell_{ij} = \frac{P_i(j \in M^\ell) \exp(V_{ij})}{\sum_{k \in J} P_i(k \in M^\ell) \exp(V_{ik})}
\]  

(3)

where \( P^\ell_{ij} \) denotes the probability we expect individual \( i \) to select brand \( j \) on a purchase occasion for usage/consumption occasion \( \ell \).

With respect to the choice probabilities shown in (3), a number of observations can be made. First, the probabilities are conditional on the usage/consumption occasion motivating the choice. This is a direct result of the usage-context dependent evaluation component underlying the choice decision. Second, other authors have postulated choice probabilities equal to those shown in (3) but did not provide a rationale or theoretical basis for doing so (see, e.g., Meyer and Cooper 1988, Fotheringham 1988). Third, the probabilities are similar in expression to those derived in other discrete choice models under completely different behavioral premises. Their structure of us-over-us-plus-them makes them compatible with Luce’s (1959) choice model as well as McFadden’s (1974) multinomial logit choice model. For \( P_i(j \in M^\ell) = 1 \) for all \( j \in J \), our choice model becomes technically a multinomial logit model. It is interesting to note that in the more general case of \( P_i(j \in M^\ell) \neq 1 \), the choice probabilities in (3) are more polarised relative to those of that nested multinomial logit model. Specifically, for brands with \( P_i(j \in M^\ell) \) close to one (i.e., the more typical brands given \( \ell \)), the choice probabilities are larger than the corresponding multinomial logit ones where the reverse is true for brands with \( P_i(j \in M^\ell) \) close to zero (i.e., the more atypical brands). Consistent with the behavioral premise, we expect typical brands to have an added choice advantage. This assertion is supported by findings in the consumer behavior literature that, at the brand level, typicality is related to the brand’s positive evaluation (Loken and Ward 1987, 1990, Nedungadi and Hutchinson 1985, and Ward and Loken 1988).

As the heterogeneity in typicality across brands is ignored in a typical multinomial logit context and, hence, all brands are considered equally substitutable, the advantage from being typical and the disadvantage from being atypical are ignored. As will be illustrated later, it is important to recognise this heterogeneity.

To further emphasize the behavioral premises of the model derived here, we contrast the model to other modeling efforts which have integrated choice set formation into discrete choice models. Manski (1977) was one of the first to discuss choice set specification in discrete choice models, and his work has motivated subsequent research in transportation
economics (see, e.g., Ben-Akiva and Boccarda 1990) and marketing (see, e.g., Roberts and Lattin 1991). Manski's (1977) argument was that since choice sets are latent, we need to define choice over all possible subsets of alternatives. Accordingly, the probability that consumer \( i \) would select brand \( j \) equals

\[
P^M_{ij} = \sum_R \Pr(C_R) \left[ \frac{\exp(V_{ij})}{\sum_{k \in C_R} \exp(V_{ik})} \right]
\]

where \( R \) is the set of all possible choice sets defined on \( J \), \( \Pr(C_R) \) is the probability that consumer \( i \) used choice set \( C_R \), and \( V_{ij} \) denotes the deterministic utility consumer \( i \) attached to brand \( j \). Evident from expression (4), the behavioral premises underlying Manski's (1977) model are that: (a) the individual consumer knows the exact utilities of the brands (i.e., utility is deterministic), (b) the consumer knows the exact content of the choice set (i.e., choice set membership is discrete and deterministic), and (c) preference order and choice on the part of the consumer are based on utility maximization. These premises are consistent with the random utility theory and the rational choice rule underlying the multinomial logit model (McFadden 1974). Note that no uncertainty arises at the individual consumer level. Modeler uncertainty is introduced in the utilities and choice set definition because of their latent character. The former is operationalised by adding a random term to the individual consumer's deterministic utility assessment, and the latter is operationalised by considering all possible subsets. This operationalisation together with the utility maximization choice rule lead to the choice probabilities shown in (4) which are essentially weighted averages of nested multinomial logit probabilities over all possible choice set subsets.

Although theoretically correct given the assumptions, the model is impractical particularly in a brand choice context where \( J \) is often very large and, hence, \( R \) is very large. Accordingly, efforts arose to reduce \( R \) without changing any of the underlying behavioral premises of the model. Ben-Akiva and Boccarda (1990) suggested the use of external constraints (e.g., brand availability) which eliminate impossible subsets from \( R \). Roberts and Lattin (1991) provided a model to reduce \( R \) to containing only one set, where the latter is defined discretely in the utility domain, using a deterministic threshold. Specifically, they argue that a brand will be in the set if and only if the consumer assesses its utility to be larger than that threshold level. Following Hauser and Wernerfelt (1990), the utility threshold level is implicitly determined by the balance of the marginal benefit of including one more brand in the choice set against the cost of considering that additional brand. In operationalising their approach, they introduce uncertainty in that utility threshold level because of its latent character on the part of the modeler. Other, more recent efforts (see, e.g., Siddarth, Bucklin, and Morrison 1993).
continue in that tradition of simplifying Manski's (1977) representation without addressing or questioning its underlying premises. Although empirical support exists for some of these modeling efforts, their weakness lies in the fact that the multinomial logit assumption of equal substitutability of brands belonging to the choice set is carried through. This is quite different from the behavioral premises underlying model (3). That the consumer will consider different brands to different degrees, conceptualised here using the goal-derived categorisation paradigm which has theoretical as well as experimental support, is addressed directly. Apart from the theoretical support to do so, recognising this element of heterogeneity and its impact on the choice task has important normative implications, as we will illustrate later. To complete the formalisation of the choice model, we first discuss the modeling of $P_j(j \in M^c)$.

**MODELING THE CHOICE SET GENERATION PROCESS**

Research in cognitive psychology (see, e.g., Rosch 1973, 1975, Rosch and Mervis 1975, Smith and Melara 1990, Fried and Holyoak 1984) and consumer behavior (see, e.g., Troye 1984, Sujan 1985, Nedungadi and Hutchinson 1985, Ward and Loken 1986, Loken and Ward 1987, Nedungadi 1990, Anderson 1991, Nedungadi and Kanetkar 1992) suggests that categories are created around exemplar or prototypical alternatives. In goal-derived categorisation (Barsalou 1983, 1985), the exemplar or prototypical alternative would be a brand which is judged to be ideal by the individual consumer given the usage/consumption goal. In forming these categories, the prototypical brand is used as a reference point (Rosch and Mervis 1975). Hence, a brand's typicality, which in our conceptualisation underlies the individual consumer's judgment about whether or not that brand belongs to the category of choice alternatives (and, hence, is a member of the choice set), arises from a relative comparison between the brand and the prototypical alternative.

In modeling this judgment, we follow consumer behavior research which suggests that the prototypical brand need not exist (i.e., be one of the brands in J); it can be an ideal profile constructed from existing brands. Furthermore, and for exactly the same reasons we specified uncertainty in brand typicality judgments, we recognise that the individual consumer can be uncertain about what is prototypical. With these characteristics, we formalise choice set generation at the individual consumer level as a rational, two-stage process. In the first stage, the individual consumer defines a set of brands which, for a specific consumption goal, are either considered typical or contribute to the prototypical profile. That particular set of brands is called the reference set and is denoted by C. In the second stage, each individual brand in set J is compared to the prototypical alternatives in C.
in order to judge its typicality and, hence, choice-set-membership likelihood, given the envisioned usage/consumption goal.

Underlying these relative evaluations is a typicality function which measures the brand’s typicality given a specific usage/consumption goal. This function consists of a deterministic component and a random component. In other words, the typicality consumer \( i \) attaches to brand \( j \) for consumption goal \( \ell \) equals

\[
T_{ij}^\ell = t_{ij}^\ell + \tau_{ij}^\ell \tag{5}
\]

where \( t_{ij}^\ell \) denotes the deterministic component of the typicality judgment, and \( \tau_{ij}^\ell \) denotes the random component. The latter component captures the consumer’s as well as the modeler’s uncertainty in the typicality judgment. The deterministic component \( t_{ij}^\ell \) contains determinants which have been identified in the categorization literature as influencing typicality. Some of these determinants could also be influencing the deterministic utility component of brand value, \( V_{ij} \) in (1). Overlap in determinants does not, however, imply dependence. Indeed, and consistent with the above discussion, we can assume that the random component \( \tau_{ij}^\ell \) in (5) is independent of the random component \( \mu_{ij} \) in (1).

**Stage 1: Formation of Reference Set C**

The brands that are considered prototypical or contributing to the prototypical profile are identified by the individual consumer using a sequential sampling process which balances typicality and evaluation cost. Behavioral support for intuitive cost/benefit calculations in such decision strategies is provided in Payne (1982), Johnson and Payne (1985), Grether and Wilde (1984), and Huber and Klein (1991). The formalisation suggests that a brand will be added (deleted) if the expected maximum typicality contained in \( n + 1 \) brands minus the same expected maximum contained in \( n \) brands exceeds (does not exceed) the additional cost of evaluating the typicality of this additional brand. Specifically, the \( (n + 1)^{th} \) brand is added as a prototypical brand if

\[
E[\max(n+1)] - E[\max(n)] > d_{n+1} \tag{6}
\]

where \( E \) is the expectations operator, \( \max(n) \) denotes the maximum typicality derived from the \( n \) prototypical brands, and \( d_{n+1} \) denotes the cost of evaluating the typicality of the \( (n+1)^{th} \) brand. Using similar arguments, dropping a brand from reference set \( C \) is a result of the reverse inequality. The resulting sampling process evolves as follows. We start with a
reference set from time \((t-1)\) of size \(n\) (where \(n < m\), \(m\) being the total number of brands in \(J\)). All other \((m-n)\) brands are candidates to enter set \(C\) at time \(t\); all \(n\) brands in set \(C\) are candidates to leave at time \(t\). The sequence in which brands will be added or dropped depends on the incremental typicality, or the left side of equation \((6)\). The implication of this procedure is that the reference set can expand and contract over time with prototypical brands entering and tending to stay in (Vanhonacker 1994).

In order to operationalize rule \((6)\), the distribution of the maximum typicality and the decision cost of evaluating the typicality of an additional brand need to be derived. Assuming that the random typicality components \(\tau_{ij}^t\) in \((5)\) are i.i.d. according to a Type-1 Extreme Value distribution, the cumulative density of the maximum given \(n\) brands equals (see, e.g. Ben-Akiva and Lerman, 1985)

\[
F_{\text{max},i}(x) = \exp\left[-\exp\left(-(x - b_i^t(n))\right)\right]
\]

which is a Type-1 Extreme Value distribution with modal value equal to

\[
b_i^t(n) = \ln \left[ \sum_{j=1}^{n} \exp(\tau_{ij}^t) \right].
\]

Given that the mean of the Type-1 Extreme Value distribution is a constant (Euler's constant) away from the modal value (Johnson and Kotz 1970), the left side of the rule in \((6)\) can be expressed as \(E[\max(n+1)] - E[\max(n)] = b_i^t(n+1) - b_i^t(n)\). Hence,

\[
E[\max(n+1)] - E[\max(n)] = \ln \left[ \sum_{j=1}^{n+1} \exp(\tau_{ij}^t) \right] - \ln \left[ \sum_{j=1}^{n} \exp(\tau_{ij}^t) \right]
\]

which, in a multinomial logit sense, equals the logarithm of the odds ratio of selecting any brand from a set of \(n\) versus selecting it from a set of \((n+1)\) brands with selection based entirely on typicality as defined in \((5)\). Note that the ratio is invariant up to an additive constant just as typicality is determined up to an additive constant.

The evaluation cost \(d_{n+1}\) in \((6)\) can be operationalized using the "cost of thinking" framework described in Shugan (1980). This operationalisation is not without criticism nor is it the only one suggested in the literature (see Payne 1982). As a comprehensive measure
of psychological inspection cost, it does provide a useful starting point. Shugan (1980) postulates that the cost of evaluating different brands against one another is directly proportional to the perceptual complexity in comparing brands, and is inversely related to both the difference in preference between the brands (see also Meyer 1982, p. 106), and the confidence at which the selection must be made. Specifically, Shugan suggests that individual i's potential cost of comparing the typicality of two brands j and k equals

\[ f_i = \frac{\text{Var}(T_{ij}^\ell) + \text{Var}(T_{ik}^\ell)}{(1-\alpha)[E(T_{ij}^\ell) - E(T_{ik}^\ell)]^2} \]

where \( \alpha \) denotes the confidence level at which the evaluation must be made and the obvious restriction that \( E(T_{ij}^\ell) \neq E(T_{ik}^\ell) \). Given the random typicality component assumptions made above, it can be shown that

\[ f_i = \frac{2\beta}{(1-\alpha)[t_{ij}^\ell - t_{ik}^\ell]^2} \]

where \( \beta \) denotes the constant variance of the Type-1 Extreme Value distribution (see Johnson and Kotz 1970, p. 278). One of the implications of the constant variance is that the scale invariance of Shugan's (1980) general results has been lost.

Following the arguments of Shugan (1980) where the evaluation cost is independent for each brand, the cost of evaluating the typicality of the additional brand can then be expressed as

\[ d_{n+1} = \frac{2\beta}{(1-\alpha)} \frac{1}{\sum_{j \neq n+1} \left( t_{ij}^\ell - t_{in+1}^\ell \right)^2} \]

Moreover, the decision to add the \((n+1)\)th brand to reference set \( C \) would depend on the inequality

\[ \ln \frac{\sum_{j=1}^{n+1} \exp(t_{ij}^\ell)}{\sum_{j=1}^{n} \exp(t_{ij}^\ell)} \frac{2\beta}{(1-\alpha)} \frac{1}{\sum_{j \neq n+1} \left( t_{ij}^\ell - t_{in+1}^\ell \right)^2} \]

(8)
Dropping the \((n+1)\) th brand already belonging to set \(C\) would depend on the reverse inequality. Hence, with the add/drop sequence as discussed above, one proceeds until a change in the set occurs (either a drop or an add depending on the inequality); at that point the incremental typicality is recalculated on the basis of the new set and a new sequence is determined. One then proceeds through the new sequence until no changes occur and the composition of the reference set at time \(t\) is defined. A numerical example illustrating the sampling process is discussed in Appendix A.

Note that the right hand side of equation (8), the evaluation cost, has a number of appealing properties. First, as the factor \((1 - \alpha)\) decreases, costs increase. As parameter \(\alpha\) can be interpreted as a measure of involvement, this suggests that lower involvement makes the reference set expand, which is consistent with Sherif and Hovland's (1964) notion that the latitude of acceptance expands with low levels of involvement. Second, the higher the difference in the deterministic component of typicality between the brands, the lower the evaluation cost. One could argue that the cost is limited because of the ease of discrimination in terms of typicality. Moreover, the directional properties of the cost components are theoretically appealing. Adaptations of (8) to ensure scale invariance, to account for inertia and learning in sequential evaluation, and to prevent infinite evaluation costs for brands with identical mean typicalities are discussed and validated in Vanhonacker (1994).

**Stage 2: Modeling \(P_j (j \in M)\)**

When a brand enters reference set \(C\), it will raise the maximum typicality contained in the set. That maximum represents a composite score of prototypicality. It is argued here that this maximum forms the reference point against which the consumer will evaluate the typicality of other brands in the universal set. In other words, the likelihood of being in the choice set is defined as the probability that the perceived typicality of a brand is larger than the maximum typicality contained in the reference set. Specifically,

\[
P_t \left( j \in M^t \right) = \text{Prob} \left( T_{ij} > b_{ij}^t \right)
\]

and, hence,

\[
P_t \left( j \in M^t \right) = \text{Prob} \left( \tau_{ij} > b_{ij}^t - t_{ij}^t \right)
\]
where $b_i^\ell$ denotes the modal value of the maximum density in (7). The modal value is used on the basis of (a) the skewed nature of the underlying distribution, (b) analytic simplicity, and (c) the fact that the mean is a constant away from the modal value. Furthermore, it makes intuitive sense to use the most likely value of the maximum as the reference point. Moreover,

$$p_i(j \in M^\ell) = 1 - \exp \left[ - \exp \left( - (b_i^\ell - t_{ij}^\ell) \right) \right]$$

(9)

with $b_i^\ell = \ln \left( \sum_{j \in C} \exp(t_{ij}^\ell) \right)$ as defined above.

This probability of belonging to the choice set has some interesting properties. The derivative of $p_i(j \in M^\ell)$ in (9) with respect to $b_i^\ell$ is clearly negative. Accordingly, anything which increases the modal value of the maximum will reduce the likelihood of belonging to the choice set for later brands. Analytic results in Appendix B show that this likelihood increases with: (a) a decrease in the number of prototypical brands already contained in the reference set, (b) a decrease in the mean and variance of the typicality judgments of the brands in that set, and (c) an increase in the uncertainty in the typicality judgments. Some interesting strategic conclusions could be drawn from these results. For example, condition (c) implies that a possible strategy for a new brand entering an established market could be to create uncertainty in or even question the brand’s typicality judgments made by consumers. In this instance, the likelihood of the new entrant being in the choice set would be enhanced.

**THE CHOICE MODEL AND ITS CHARACTERISTICS**

Integrating the operationalization of the choice set generation process into the choice model shown in (3) implies that the probability that consumer $i$ will choose brand $j$ for usage/consumption occasion $\ell$ becomes

$$p_{ij}^\ell = \frac{\left[ \exp(V_{ij}) \right] \left[ 1 - \exp \left( - \exp \left( - (b_i^\ell - t_{ij}^\ell) \right) \right) \right]}{\sum_{k \in J} \left( \exp(V_{ik}) \right) \left[ 1 - \exp \left( - \exp \left( - (b_i^\ell - t_{ik}^\ell) \right) \right) \right]} \quad \text{for } j \in J.$$  

(10)

18
This choice model has some interesting and intuitively appealing characteristics. First, the model does not contain the unappealing Independence of Irrelevant Alternatives (IIA) property which characterizes many choice models. This property implies that when a new brand is introduced, it will obtain market share proportionally from all brands regardless of substitutability. In the above model, the new brand will obtain proportionally more market share from the prototypical brands (i.e., the brands belonging to reference set C) regardless of whether or not it is similar to them. The latter is in line with asymmetric dominance which was shown to exist in Huber, Payne, and Puto (1982).

In the discussion above, we contrasted model (10) to the nested multinomial logit model (2) which contains the brand value determinants. However, it is interesting at this point to also contrast model (10) to a multinomial logit model containing all determinants (i.e., the brand value as well as the typicality determinants). Indeed, one might argue that specifying a random utility model containing both $V_{ij}$ in (1) and $t_{ij}^{\ell}$ in (5) together with the utility maximizing choice rule would provide a strong competitor to model (10), particularly given the known robustness of the logit model. Following McFadden (1974), that model would predict individual consumer i's choice probability for brand j to equal

$$p_{ij}^{MNL} = \frac{\exp(V_{ij} + t_{ij}^{\ell})}{\sum_{k \in J} \exp(V_{ik} + t_{ik}^{\ell})} \quad \text{for } j \in J.$$  

(11)

As with the probabilities in (10), these logit probabilities are also conditional on $\ell$, the usage/consumption occasion motivating the choice. This conditional nature is typically not recognised in the development and application of the multinomial logit model, and constitutes one point of departure. Another important difference are the behavioral premises underlying the models. The logit model was developed in the framework of traditional random utility theory (Manski 1977) and assumes that (a) there is no uncertainty on the part of the individual consumer, (b) the choice set equals J with all brands being equally substitutable (Ben-Akiva and Lerman 1985), and (c) the choice rule is one of utility maximization. These assumptions are quite different from the ones underlying choice model (10). However, the multinomial logit model can be derived within the conceptualisation of model (10) If, for example $P_{i}(j \in M^{\ell}) = \exp(t_{ij}^{\ell}) / \sum_{k \in J} \exp(t_{ik}^{\ell})$, then the choice rule of jointly maximizing $P_{i}(j \in M^{\ell})$ and $P_{i}(j)$ adopted here leads to (11). Such a likelihood of choice set inclusion would arise when the individual consumer would maximize typicality across all brands in J. With the distributional assumptions on $t_{ij}^{\ell}$ in (5) as specified above, maximizing typicality across all $j \in J$ results in a logit formulation of the probabilities of
choice set membership. Although technically identical, the resulting choice model is conceptually very different from the traditional multinomial logit model. Furthermore, in contrast to model (10), it implies a very different choice set generation process. Maximizing typically implies that in creating the choice set, the individual consumer makes pairwise typicality judgments on all brands available. There is no support in the categorisation literature for such a process; the theoretical arguments as well as experimental support discussed above suggest that brands are compared individually to an exemplar (or prototypical) brand. The latter process is precisely the one underlying (10). In other words, cognitive psychology and consumer behavior research support the choice set generation process underlying model (10) and not the one underlying (11). Nevertheless, and particularly given the known robustness of the logit model, model (11) is a good null model for validation purposes. Furthermore, comparing (11) to (10) provides an empirical test of the different choice set generation processes within the adopted conceptualisation.

EMPIRICAL ILLUSTRATION

The brand choice model derived and shown in (10) was validated empirically for a single product category. For this illustration, supermarket scanner data were made available by A.C. Nielsen-France from their SCAN-7000 panel together with telephone-based survey data on brand awareness among the panel members. The product category selected is a frequently-purchased food item used in France almost exclusively as a bread spread. This characteristic enabled us to treat all purchases as essentially being motivated by a single and identical usage/consumption goal. The validation discussed here is confined to the choice set formation process and the choice model; the available data do not enable a direct validation of the reference set formation process as formalised above. We do, however, rely on external data for defining reference set C which are consistent with that process.

The product category studied has three manufacturers marketing seven brand names which compete with store brands. The brands come in three different quality varieties with not all brand names marketing all varieties. All brands are available in a large number of flavors. In this analysis, we do not distinguish flavors and focus on the eight brands (i.e., all outlet-specific store brands are lumped together in a single "store" brand).

The sample available consisted of 210 families who each made at least 12 purchases in the category over a 65 week period for which data were available. In all, 4972 purchase records were included in the analysis. For each household, the historical purchase records were broken down in three, approximately equal, subsamples of purchases. The first subsample, containing 1589 purchase records, was used for initialisation purposes. The second
subsample, containing 1794 purchase records, was used for the estimation, and the third subsample, containing 1589 purchase records, was used for the validation.

For each purchase record, the following information was available: the brand purchased, the price paid, which other brands were available in the store at the time of purchase, shelfspace allocated to each brand in the store at the time of the purchase (measured in meters), and which brands were on promotion (measured as a dummy variable). Promotions in this category are primarily temporary deals providing an immediate monetary benefit (e.g. cents off, two-for-the-price-of-one, etc.). These data were supplemented with brand awareness data. During the time period covered, the person in the household who does the shopping was called and asked about his/her awareness of different brand names. The awareness data covered the seven individual brand names as well as the store brand. With respect to the latter, no data were collected on the individual store brands; awareness of a specific store brand was coded simply as awareness for the "store" brand. The coding system used was as follows: 3 for top-of-the-mind unaided awareness (i.e., brand mentioned first), 2 for unaided awareness (i.e., brand mentioned later), 1 for aided awareness (i.e., the family member recognised the brand name when given), and zero otherwise (i.e., unaware of brand even after been given the name). A summary of the data for both the estimation sample and the validation sample is given in Table 1.

The sample statistics by themselves already reveal some interesting patterns supporting the structure of the choice model derived. Compare, for example, the statistics for the store brand and brand 3. The store brand has by far the largest purchase share (i.e., 41% of the purchases in the estimation sample). At the same time, its awareness is far below that of some other brands and about half that of brand 3. The average awareness score for the store brand is particularly low considering its aggregate nature. As discussed above, recalling or recognising only one of the many store brands would be counted as being aware. Apparently, store brand awareness (both aided and unaided) is particularly low relative to brand name awareness in this product category. Brand 3, on the other hand, has the highest awareness among the 8 brands and has the second highest market share. The store brand, as one would expect, has the highest promotion frequency (more than twice that of brand 3) and a relatively low price (on average 16.49 versus 19.93 for brand 3). Accordingly, where brand 3 seems to be getting share points from awareness, the store brand seems to capitalise on the monetary substitution value it provides. Whether this interpretation is correct needs further investigation, but the pattern supports the argument for considering and distinguishing the two brand evaluation components underlying choice as discussed above.

Brand typicality $t_{ij}$ in (5) was specified as a linear additive function of three variables: loyalty, awareness, and shelfspace. Loyalty, measured here as suggested in Guadagni and
<table>
<thead>
<tr>
<th></th>
<th>Purchase Set&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Availability&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Promotion</th>
<th>Shelfspace</th>
<th>Price Paid</th>
<th>Loyalty&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Awareness</th>
<th>Purchase Share</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Estimation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brand 1</td>
<td>0.32</td>
<td>0.92</td>
<td>0.052</td>
<td>2.03</td>
<td>18.14</td>
<td>0.105</td>
<td>1.391</td>
<td>0.111</td>
</tr>
<tr>
<td>Brand 2</td>
<td>0.17</td>
<td>0.35</td>
<td>0.075</td>
<td>1.84</td>
<td>17.40</td>
<td>0.039</td>
<td>0.822</td>
<td>0.033</td>
</tr>
<tr>
<td>Brand 3</td>
<td>0.55</td>
<td>0.99</td>
<td>0.116</td>
<td>4.09</td>
<td>19.93</td>
<td>0.239</td>
<td>1.936</td>
<td>0.249</td>
</tr>
<tr>
<td>Brand 4</td>
<td>0.20</td>
<td>0.88</td>
<td>0.034</td>
<td>1.27</td>
<td>21.07</td>
<td>0.087</td>
<td>1.028</td>
<td>0.081</td>
</tr>
<tr>
<td><strong>Sample</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brand 5</td>
<td>0.14</td>
<td>0.42</td>
<td>0.012</td>
<td>0.48</td>
<td>18.50</td>
<td>0.030</td>
<td>0.513</td>
<td>0.027</td>
</tr>
<tr>
<td>Brand 6</td>
<td>0.11</td>
<td>0.30</td>
<td>0.000</td>
<td>0.79</td>
<td>18.78</td>
<td>0.029</td>
<td>0.893</td>
<td>0.025</td>
</tr>
<tr>
<td>Brand 7</td>
<td>0.16</td>
<td>0.40</td>
<td>0.067</td>
<td>0.68</td>
<td>16.36</td>
<td>0.067</td>
<td>1.509</td>
<td>0.065</td>
</tr>
<tr>
<td>Store Brand</td>
<td>0.76</td>
<td>1.00</td>
<td>0.287</td>
<td>7.79</td>
<td>16.49</td>
<td>0.409</td>
<td>1.075</td>
<td>0.410</td>
</tr>
<tr>
<td><strong>Validation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brand 1</td>
<td>0.38</td>
<td>0.86</td>
<td>0.028</td>
<td>2.02</td>
<td>19.28</td>
<td>0.084</td>
<td>1.385</td>
<td>0.082</td>
</tr>
<tr>
<td>Brand 2</td>
<td>0.22</td>
<td>0.32</td>
<td>0.098</td>
<td>1.86</td>
<td>18.27</td>
<td>0.040</td>
<td>0.818</td>
<td>0.042</td>
</tr>
<tr>
<td>Brand 3</td>
<td>0.65</td>
<td>0.99</td>
<td>0.106</td>
<td>4.09</td>
<td>20.07</td>
<td>0.264</td>
<td>1.928</td>
<td>0.271</td>
</tr>
<tr>
<td>Brand 4</td>
<td>0.28</td>
<td>0.94</td>
<td>0.054</td>
<td>1.30</td>
<td>22.14</td>
<td>0.075</td>
<td>1.029</td>
<td>0.069</td>
</tr>
<tr>
<td><strong>Sample</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brand 5</td>
<td>0.17</td>
<td>0.53</td>
<td>0.007</td>
<td>0.54</td>
<td>19.57</td>
<td>0.026</td>
<td>0.518</td>
<td>0.026</td>
</tr>
<tr>
<td>Brand 6</td>
<td>0.12</td>
<td>0.33</td>
<td>0.000</td>
<td>0.49</td>
<td>19.42</td>
<td>0.021</td>
<td>0.907</td>
<td>0.018</td>
</tr>
<tr>
<td>Brand 7</td>
<td>0.23</td>
<td>0.53</td>
<td>0.032</td>
<td>0.96</td>
<td>16.65</td>
<td>0.069</td>
<td>1.512</td>
<td>0.071</td>
</tr>
<tr>
<td>Store Brand</td>
<td>0.82</td>
<td>1.00</td>
<td>0.172</td>
<td>7.89</td>
<td>17.19</td>
<td>0.416</td>
<td>1.073</td>
<td>0.420</td>
</tr>
</tbody>
</table>

<sup>a</sup> Proportion of purchases the brand was in the historical purchase set (i.e., a prior purchase of the brand was recorded).
<sup>b</sup> Proportion of purchases the brand was available.
<sup>c</sup> Loyalty measure as defined in Guadagni and Little (1983).
Little (1983), captures repeat purchase tendencies as a result of either inertia or positive preference evaluations. Both familiarity and preference have been linked to typicality in the consumer behavior literature (see, e.g., Zajonic 1968, Loken and Ward 1990), and we expect the loyalty variable to capture the positive effect these determinants have on brand typicality. Furthermore, and as already mentioned, some consumer behavior research suggests that there is little external information search in grocery shopping; most information used in brand evaluations is recalled from memory (see, e.g., Hoyer (1984) and Dickson and Sawyer (1986)). In addition, research has established that memory recall is aided by external cues (see, e.g., Alba et al. 1991). We argue that allocated shelfspace is an important cue in aiding recall and recognition. Furthermore, and particularly so for consumers less familiar with the category, allocated shelfspace might be interpreted more directly and in a more general context as a signal of the brand's typicality. Accordingly, we expect shelfspace to have a positive impact on brand typicality. Finally, and consistent with the consumer behavior research finding that prototypical brands have a memory-based advantage (see, e.g., Alba et al. 1991), we expect awareness to be positively related to brand typicality. In essence, loyalty, shelfspace, and awareness capture the actual behavior characteristics as well as the memory retrieval characteristics which have been linked to brand typicality in the consumer behavior literature.

The brand value component \( V_{ij} \) in (1), was specified as a linear additive function of: intrinsic value, price paid, and promotion. Consistent with Lattin and Bucklin (1989), brand value is defined in the price-preference domain. The intrinsic value is the traditional brand-specific constant specified in discrete choice models which captures the individual consumer's intrinsic preference for the brand. Note that these preferences are specified explicitly in \( V_{ij} \) and are contained implicitly in the loyalty variable specified in \( t_{ij}^l \). In addition to the intrinsic value, we specify the monetary value of the offer. Both price and promotion are specified as determinants of that value. We expect price to have a negative impact and promotion to have a positive impact on brand value.

So far, we have operationalized the functional form and the predictors of both \( t_{ij}^l \) in (5) and \( V_{ij} \) in (1). The only variable which remains to be operationalized is \( C \) or the reference set. As discussed, \( C \) contains the brands which jointly represent the reference point (or prototypical profile) used in the typicality judgments. Using arguments similar to those justifying loyalty as a determinant of typicality, we specify \( C \) as the historical purchase set. In other words, for each choice occasion the assumption was made that the prototypical profile was derived from the brands the household had purchased up to that point. This particular definition might be somewhat narrow as one could argue that brand awareness and knowledge and, hence, the construction of the prototypical reference brand are not
necessarily limited to the historical purchase set. Indeed, the reference set formation process discussed above does not imply that the incremental benefit would outweigh evaluation costs only for brands purchased.

With the operationalization of reference set C and the determinants of $t_{ij}^c$ and $V_{ij}$, the choice model in (10) is fully operationalised. The model was estimated using maximum likelihood across all households. An overview of the estimation results and some validation statistics based on the holdout sample of purchases is given in Table 2. For comparison purchases, we show the estimation and validation results for a nested version of the proposed model (10) where the awareness variable was suppressed in the typicality specification, and for the multinomial logit model (11) with and without the awareness variable. Comparing the results for model (10) with its nested version provides a test of the contribution of brand awareness which is featured as a key variable in the usage-context dependent brand evaluations underlying choice set formation. The multinomial logit model without the awareness variable is used as a benchmark representing the more traditional specification of the model with all determinants contained in the deterministic utility component. Consistent with the random utility assumptions underlying that model, it would be difficult to argue that awareness should be included as providing additional (and independent) utility value. The multinomial logit model with awareness represents the null model discussed above and shown in (11). Comparing that model with model (10) provides an empirical test of the different choice set generation processes as discussed above. We first discuss the parameter estimates to familiarise ourselves with the substantive response characteristics of this product category.

In general, the parameter estimates shown in Table 2 have the correct sign. An exception is price but this predictor is highly insignificant in all four models. The latter could be attributed to the large variance in prices across flavors under the same brand name which is not directly accounted for here. Apart from price, the promotion effect is positive as expected and highly significant. Hence, promotion is identified as a strong predictor of brand value. Focusing on the typicality predictors, shelfspace, awareness, and loyalty have a positive impact as expected. Where the latter two are highly significant, shelfspace is only marginally significant. The strong effect of loyalty (and, hence, the purchase history) indirectly supports the operational definition of reference set as the historical purchase set. Note that the results imply that typicality in this category is determined by awareness and past purchase history.

Focusing on the various validation statistics computed on the holdout sample and shown in Table 2, model (10) does quite well in absolute and relative terms. Using the likelihood ratio test, model (10) is better than its nested alternative at the 0.01 significance level (chi-square
Table 2
Estimation and Validation Results

<table>
<thead>
<tr>
<th></th>
<th>Multinomial Logit Model (11)</th>
<th>Proposed Model (10)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Without Awareness</td>
<td>With Awareness</td>
</tr>
<tr>
<td>Typicality Predictors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loyalty</td>
<td>3.722 (43.78)(^a)</td>
<td>3.441 (33.94)</td>
</tr>
<tr>
<td>Shelfspace</td>
<td>0.022 (1.80)</td>
<td>0.011 (0.93)</td>
</tr>
<tr>
<td>Awareness</td>
<td>-</td>
<td>0.270 (6.76)</td>
</tr>
<tr>
<td>Value Predictors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brand Constants</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-0.255 (-1.27)</td>
<td>-0.145 (-0.74)</td>
</tr>
<tr>
<td>2</td>
<td>0.090 (0.77)</td>
<td>-0.066 (0.56)</td>
</tr>
<tr>
<td>3</td>
<td>-0.334 (-2.21)</td>
<td>-0.334 (-2.23)</td>
</tr>
<tr>
<td>4</td>
<td>-0.378 (-1.98)</td>
<td>-0.144 (-0.76)</td>
</tr>
<tr>
<td>5</td>
<td>-0.570 (-2.25)</td>
<td>-0.432 (-1.99)</td>
</tr>
<tr>
<td>6</td>
<td>-0.059 (-0.32)</td>
<td>-0.201 (-1.13)</td>
</tr>
<tr>
<td>7</td>
<td>0.013 (0.10)</td>
<td>0.242 (1.73)</td>
</tr>
<tr>
<td>Price Paid</td>
<td>0.014 (0.77)</td>
<td>0.020 (1.16)</td>
</tr>
<tr>
<td>Promotion</td>
<td>0.258 (3.51)</td>
<td>0.281 (3.87)</td>
</tr>
<tr>
<td>Number of Parameters</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Validation(^b)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R^2)-Statistic</td>
<td>0.508</td>
<td>0.510</td>
</tr>
<tr>
<td>Hitrate</td>
<td>0.722</td>
<td>0.726</td>
</tr>
<tr>
<td>Mean (Variance) of Predicted Probability for Chosen Brand</td>
<td>0.614 (0.111)</td>
<td>0.616 (0.112)</td>
</tr>
</tbody>
</table>

\(^a\) Asymptotic t-value

\(^b\) Holdout-sample validation results
value of 36 with 1 degree of freedom). Accordingly, as a predictor of typicality awareness provides a significant contribution. Given that it is theoretically difficult to argue that awareness of a brand should be specified as an additive component in that brand's utility function, this result casts some doubt on the usefulness of the traditional multinomial logit model to explain choice behavior in this category.

Relative to the multinomial logit results, the proposed model does better in terms of $U^2$-statistic as well as level of mean predicted probability for the brand chosen. On the hitrate in the holdout sample, model (10) does better than the nested multinomial logit model but equally well than the fully specified multinomial logit model (11). Although direct statistical tests such as the likelihood ratio test are not possible, note that the absolute improvement in $U^2$ and mean purchase probability for the brand chosen provided by model (10) over both multinomial logit models is larger than the corresponding improvement over its nested version which was conclusively rejected as inferior. This is even the case for the fully specified multinomial logit model with one parameter more than the nested version of model (10). Hence, we can conclude with some confidence that model (10) and its underlying premises explains choice behavior better for this particular product category. This conclusion will be further supported with some brand-level insights discussed next.

Underlying the aggregate results reported in Table 2 are some insightful patterns at the brand level. Not only do they provide a better understanding of the various brands in this category, they further illustrate the importance of recognising the behavioral characteristics of the adopted conceptual framework. The brand-level insights provided by model (10) are particularly useful to a brand manager who is trying to understand his/her brand. Relying on the brand-level choice probabilities, we derive normative recommendations for promotion frequencies to illustrate their characteristics and interpretability in the context of the premises of model (10) and contrast them to those derived from the multinomial logit model. The latter illustrate that not fully recognising the heterogeneity in typicality across brands and its origin as captured and operationalised in (10) leads possibly to erroneous insights as well as suboptimal behavior.

Table 3 provides an overview of the choice predictions at the brand level for the estimation sample. For each brand, the table contains the choice set membership probabilities and choice probabilities derived from proposed model (10) and multinomial logit model (11). As can be seen, the corresponding probabilities are quite different. Model (10) predicts a probability of choice for the store brand which is much higher than the probability predicted by the multinomial logit model. In contrast, the reverse is true for the other name brands. There, model (10) predicts lower choice probabilities than the multinomial logit model. A
Table 3
Brand Specific Estimation Results\textsuperscript{a}

<table>
<thead>
<tr>
<th></th>
<th>Multinomial Logit Model (11)</th>
<th>Proposed Model (10)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(P_1(j \in M^\ell)^b)</td>
<td>Purchase Probability</td>
</tr>
<tr>
<td>Brand 1</td>
<td>0.122</td>
<td>0.1268</td>
</tr>
<tr>
<td>2</td>
<td>0.061</td>
<td>0.0491</td>
</tr>
<tr>
<td>3</td>
<td>0.227</td>
<td>0.2680</td>
</tr>
<tr>
<td>4</td>
<td>0.095</td>
<td>0.0731</td>
</tr>
<tr>
<td>5</td>
<td>0.039</td>
<td>0.0278</td>
</tr>
<tr>
<td>6</td>
<td>0.055</td>
<td>0.0322</td>
</tr>
<tr>
<td>7</td>
<td>0.101</td>
<td>0.0970</td>
</tr>
<tr>
<td>Store Brand</td>
<td>0.299</td>
<td>0.3261</td>
</tr>
</tbody>
</table>

\textsuperscript{a} Evaluated on the estimation sample.

\textsuperscript{b} \(P_1(j \in M^\ell) = \exp(t_{ij}^\ell) / \sum_{k \in J} \exp(t_{ik}^\ell)\).

\textsuperscript{c} \(P_1(j \in M^\ell) = 1 - \exp[-\exp(-(b_i^\ell - t_{ij}^\ell))]\) or expression (9).
similar pattern is present in the choice set membership probabilities. Hence, a clear
difference emerges in how both models treat the store brand versus the name brands.

Interestingly enough, in almost all cases, the corresponding purchase share figures shown in
Table 1 are in between the predictions of both models. Specifically, where model (10)
overpredicts the store brand's choice share, the multinomial logit model underpredicts it.
The reverse is true for the other 7 name brands. These results arise from the fact that the
multinomial logit model does not adequately capture brand typicality in this product
category. As discussed above already, brand typicality and its heterogeneity across brands
leads to relatively more polarised choice probabilities for model (10). This is evident in the
results shown in Table 3. For example, the multinomial logit model implies choice set
membership probabilities for the store brand and brand 3 which are not that different (0.299
versus 0.227, respectively). This despite the store brand's significantly higher promotional
frequency (0.287 versus 0.116, respectively, in the estimation sample) and purchase share
(0.410 versus 0.249, respectively, in the estimation sample). Furthermore, the estimation
results in Table 2 suggested strong response to promotional activity in the category. Hence,
sustained relatively high levels of promotional activity by the store brand should help
maintain a high purchase frequency over repeated choice occasions. The latter, given loyalty
(operationalised on purchase history) being identified as a strong predictor of typicality
besides awareness, should lead to a relatively high choice set membership. The latter comes
out in the model (10) results but not in the multinomial logit model results. Given the
relatively strong awareness for brand 3 as discussed above (and shown in Table 1), the
multinomial logit model seems to implicitly give almost equal weight to awareness and
loyalty in the choice set membership probabilities. As discussed in the conceptualisation of
choice set generation underlying model (10), the store brand would be the reference point in
the typicality judgments. Accordingly, the profile of that brand on the typicality
determinants identified in this category defines the relative weighing scheme of loyalty
versus awareness in assessing choice set membership. A different weighing scheme might
arise, and apparently does considering the results, in the typicality maximizing logit
formulation underlying the multinomial logit model as discussed above. Where brand 3's
choice set membership probability implied by the logit model is not that different from the
one in model (10) (0.227 versus 0.229, respectively), the store brand's choice set
membership probability is very different with model (10) predicting it to be 40% higher
(0.299 versus 0.421, respectively). This difference underlies the disparity in predicted
choice probabilities. Although neither model predicts the choice probabilities exactly, the
conceptualisation of brand choice adopted here enables an insightful and intuitively
appealing interpretation of the results. The results strongly support the separation of brand
value and choice set membership. With respect to the latter, the underlying heterogeneity in
typicality across brands seems to be very important. The results do indicate, however, that
the operationalisation of choice set generation incorporated in (10) somewhat overstates that heterogeneity and improvements in the modeling of that component can be made. We proceed with discussing the normative implications of model (10) which further illustrates the insightfulness provided by the adopted conceptualisation of brand choice.

In subsequent analysis, we derive the optimal promotion frequencies for the eight brands and illustrate the disparity in recommendation between the multinomial logit model results and the new model results. We illustrate this in a simple, static, profit maximizing framework given atomistic competition and assuming that the market is not expanding. We believe that relaxing these assumptions will not alter the relative pattern of deviation in normative recommendations given by both models.

Under the stated conditions, the optimal promotion frequency for brand \( j \) equals

\[
\text{Promo}_j^* = \left( MS_j \cdot M \cdot MG_j \cdot e_{\text{promo}} / FC_{\text{promo}} \right)
\]

where \( MS_j \) denotes the market share of brand \( j \), \( M \) denotes the size of the market (hence, \( MS_j \cdot M \) equals brand \( j \)'s total sales volume), \( MG_j \) is brand \( j \)'s gross margin per unit, \( e_{\text{promo}} \) denotes brand \( j \)'s promotion elasticity, and \( FC_{\text{promo}} \) is a factor which relates the promotional frequency to fixed promotional cost for brand \( j \) (i.e., \( FC_{\text{promo}} \cdot \text{Promo} \) equals the total fixed costs of promotion). This normative result is derived straightforwardly from the static profit model

\[
PR_j = ( MG_j \cdot M \cdot MS_j ) - ( FC_{\text{promo}} \cdot \text{Promo} )
\]

where \( MS_j \) is a function of \( \text{Promo} \). Accordingly, and substituting \( MS_j \) by the predicted choice probabilities, comparing the optimal promotion frequencies recommended by the multinomial logit model and the proposed model comes down to comparing their respective promotion elasticity estimates times their choice probability predictions. By definition, the promotion elasticity of either model equals

\[
e_{\text{promo}} = [ \beta_{\text{promo}} \cdot \text{Promo} ] \cdot (1 - MS_j)
\]

where \( \beta_{\text{promo}} \) is the estimated promotion response parameter. Accordingly, the relative comparison in normative predictions comes down to evaluating

\[
[ MS_j \cdot \beta_{\text{promo}} \cdot (1 - MS_j) ]
\]

for both models. Substituting the respective estimates for \( \beta_{\text{promo}} \) and the derived choice probability estimates, we can derive the ratio of optimal promotion frequencies recommended by the multinomial logit model over those recommended by the proposed model (10). If the ratio is larger than 1, model (10) recommends a lower frequency than the multinomial logit model; if that ratio is smaller than 1, the reverse is true. The empirical results are shown in Table 4. The brands are listed in order of choice set inclusion probability.
Table 4
Comparison of Optimal Promotion Frequencies Recommended by the Multinomial Logit Model and the Proposed Model

<table>
<thead>
<tr>
<th>Actual Promotion Frequency</th>
<th>Ratio(c)</th>
<th>Actual Brand Awareness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Store Brand(^b)</td>
<td>0.287</td>
<td>0.896</td>
</tr>
<tr>
<td>Brand 3</td>
<td>0.116</td>
<td>1.196</td>
</tr>
<tr>
<td>Brand 1</td>
<td>0.052</td>
<td>1.525</td>
</tr>
<tr>
<td>Brand 7</td>
<td>0.067</td>
<td>1.629</td>
</tr>
<tr>
<td>Brand 4</td>
<td>0.034</td>
<td>1.355</td>
</tr>
<tr>
<td>Brand 6</td>
<td>0.000</td>
<td>1.294</td>
</tr>
<tr>
<td>Brand 2</td>
<td>0.072</td>
<td>1.523</td>
</tr>
<tr>
<td>Brand 5</td>
<td>0.012</td>
<td>1.134</td>
</tr>
</tbody>
</table>

\(a\) Evaluated for the estimation sample.

\(b\) Brands are ordered in terms of size of choice set inclusion probability.

\(c\) If larger than one, the multinomial logit model (11) recommends a higher promotion frequency than model (10); the reverse is true when the ratio is less than one.
Not surprisingly given the above discussion, the ratios shown in Table 4 indicate quite a significant departure in normative predictions. For brand 7, for example, the multinomial logit model recommends an optimal promotion frequency which is 62.9% higher than the frequency recommended by model (10). Looking across the 8 brands, a pattern emerges in the disparities. For the more typical store brand, the proposed model (10) suggests that it can enhance its profitability by increasing its already high level of promotional frequency beyond the level recommended by the multinomial logit model. In contrast, for all other (less typical) brands, model (10) suggests that in terms of profitability, they are better off reducing their promotional frequencies relative to the level recommended by the multinomial logit model. Furthermore, for the more aware brands (i.e., brands 3, 1, and 7), the disparity in optimal promotion frequency recommendations increases with reduced brand typicality (and relative share position).

The results suggest that the store brand is quite different from the name brands in terms of building market share. The store brand builds share on promotional activity and, hence, the brand value component underlying the choice decision. With repeated purchases affecting its typicality through the loyalty variable, the store brand maintains its likelihood of being in the choice set despite very low levels of awareness. Model (10) seems to recognise this unique position and recommends a further strengthening of that position through increased promotional activity beyond the level recommended by the multinomial logit model. In contrast, the name brands seem to find the basis of their share position in their typicality as determined by awareness. Relative value created through promotional activity adds but, as the normative recommendations suggest, that support is less critical and can be reduced relative to the multinomial logit model recommendations. The latter is particularly the case if the brand cannot maintain its awareness and, hence, its likelihood of being in the choice set. One insight which seems to emerge is that the name brands might be better off building awareness (and, as such, increasing their likelihood of being in the choice set) where the store brand is better off providing substitution value (and through increased frequency of choice over time increase its likelihood of being in the choice set). Although an intuitively appealing and intriguing proposition, more research is needed to corroborate this assertion. Nevertheless, as the example and the discussion illustrates, the potentially useful managerial and theoretical insights provide compelling support for the brand choice conceptualisation adopted in this research.

CONCLUSION AND FUTURE RESEARCH

In this paper, an individual-level choice model was developed which integrates a choice set formation process into the brand selection task. Choice set formation is conceptualized as a
goal-derived categorization process and is operationalised as a rational, two-stage process based on individual-level judgments about the typicality of each brand for a specific usage/consumption goal which motivates the choice. Uncertainty surrounding the latter gives rise to a probabilistic definition of choice set membership on the part of the individual consumer. In selecting a brand, the individual consumer is assumed to maximize jointly the probability of the brand being included in the choice set and the brand's relative value where the latter is determined independently of any usage/consumption occasion. A parsimonious representation of the choice probabilities is derived recognising the latent character of both evaluation components.

The modeling framework brings together a number of different areas. The categorization paradigm and the typicality construct find strong support in the cognitive psychology literature. Recently, the consumer behavior literature has further expanded that work primarily looking at typicality and its determinants at the brand level. Most of this work has been conceptual, however, and this paper provides an initial but theoretically motivated attempt to operationalize the concept in the context of brand choice. Another contribution to the literature on categorization and typicality is the explicit recognition of an element of uncertainty underlying these constructs so far not recognized in the consumer behavior literature.

Choice set generation as goal-derived categorization fits into the current stream of research in consumer behavior which claims that all purchase behavior is aimed to achieve certain goals. This paper provides a formalisation of that paradigm within a discrete choice framework. In contrast to all previous discrete choice models, the choice model developed here recognizes that the individual consumer will attach uncertainty to the alternatives in the choice set. Within the conceptual framework developed, an alternative interpretation is provided for the familiar multinomial logit model of choice.

The choice probabilities derived in this research are conditional on the usage/consumption occasion motivating the choice decision. The conditionality arises from the conceptualisation that choice sets are constructed in function of a usage/consumption goal. Under the assumptions about brand typicality postulated here, the model is conditional on the consumption goal but it is not limited to that. Indeed, evidence exists that typicality is determined by versatility. Ratneshwar and Shocker (1991) have demonstrated in a goal-derived categorization framework that typicality is strongly related to versatility defined as the number of usages for which a brand is perceived appropriate. Accordingly, a brand which is viewed as being helpful to achieve various consumption or usage goals would be viewed as more typical than one that is viewed as being helpful in achieving one single goal. Note that this would imply that the reference point (the prototypical item) used in typicality

32
judgments would be constant across consumption or usage goals. Accordingly, neither \( P_l(j \in M_l^c) \) nor the choice probabilities \( P_{ij}^c \) would be conditional on the consumption goal.

An empirical illustration of the choice model for a single product category supports the model. Relative to the multinomial logit model, the proposed model describes brand choice better according to various validation criteria. The illustration also supports the determinant role of awareness in brand typicality judgments (and, hence, choice set membership). An investigation of the brand-level predictions provides interesting insights into store brands versus name brands. For the product category investigated, the store brands seem to generate their business - and support their central position in the category - from the extensive promotions they run on top of low pricing. Hence, the monetary substitution value, which is usage-context independent, provides the basis for their share positions. In contrast, the name brands seem to generate their business from determinants directly affecting choice set membership such as awareness. Hence, they need to be considered in the context of the usage/consumption occasion motivating the choice to have a chance. Furthermore, an insightful picture and discussion is provided of the normative recommendations implied by the proposed choice model relative to those implied by the multinomial logit model. The results indicate that the multinomial logit model estimates understate the profitability of promotional activities by the store brand (the more typical brand in the context of the empirical results) and overstates the profitability of those activities for the name brands. Although not investigated further here, the results seem to suggest that the name brands might be better of concentrating on their awareness (and, hence, activities directly impacting on brand awareness) to enhance their likelihood of being considered as a choice alternative. Given its managerial relevance, further research on this assertion is warranted.

Although the illustration and discussion supports the conceptual framework and the consideration and separation of the two brand evaluation components underlying a brand choice decision, the brand-level choice probability predictions suggest that further research is warranted on the modeling of choice set membership within the behavioral framework suggested. In contrast to the traditional multinomial logit model, the operationalisation suggested here does address the heterogeneity in brand typicality and, hence, unequal substitutability across brands. However, in the product category investigated the model seems to overpredict the choice set membership of the more typical store brand and underpredict the choice set membership of the more atypical name brands. In other words, the polarisation of the membership probabilities (and, hence, choice probabilities) is too strong. Indeed, the empirical results indicate that the correct probabilities are somewhere between the logit membership probabilities (underlying (11)) and the ones implied by the
suggested model (underlying (10)). Research should be devoted to enhancing choice set membership predications in that direction.

The choice model opens up other interesting avenues for further research, particularly in the direction of pertinent managerial questions such as order-of-entry (or pioneering) advantages in emerging product categories and successful brand extensions. The concept of typicality has featured centrally in some innovative work on pioneering advantages (see, Carpenter and Nakamoto 1989). The model developed in this research provides an integrative analytic framework to possibly study whether or not these advantages exist and if they do, what gives rise to them. The most important application of the typicality construct in the consumer behavior literature has been its use in explaining and predicting the success of brand extensions (see, e.g., Herr et. al. 1990, Boush and Loken 1991, Dawar and Anderson 1993). Again, the choice model framework developed in this paper provides a new vehicle to study questions in this important area.

Besides further theoretical and conceptual developments linked to the typicality construct and the choice set generation process, the choice model itself provides a number of challenges. First of all, more extensive and comprehensive empirical work is needed. More work is also needed on the reference set generation process. The operationalisation provided in this paper was not directly validated because data were not available to do so. Experimental research might be called for to improve this component of the model.

Another direction of future research is the role of uncertainty in the typicality judgments. The choice rule adopted here implies risk avoidance on the part of the individual consumer. However, as extensive research on decisions under uncertainty suggests, attitudes towards risk which are different across consumers are important to be taken into account.

Finally, the insights discussed in the conceptual development of the model and the empirical illustration point to a fruitful avenue for managerially-useful normative work. The separation of choice set generation and brand selection implies a heterogeneity with respect to the consumer's response to external marketing factors. More conceptual and empirical work on either process will enable us to further characterize the impact of those external marketing factors and derive more optimal decision rules for them. The behavioral richness of the model provides potentially useful information for manufacturers on more efficient and effective marketing programs.
REFERENCES


Appendix A: Numerical Example of Reference Set Formation Process

Assume the deterministic typicality values for 5 brands ( \( m = 5 \)) at time \( t \) are,

\[
\begin{align*}
\ell(t) &= 6 & \ell(t) &= 8 & \ell(t) &= 5 \\
2 & & 5 & & 5
\end{align*}
\]

Furthermore, the composition of the reference set at \( t - 1 \) is: 1 0 0 1 0, or \( \{ 1, 4 \} \) (and, hence, \( n = 2 \)), and \( \alpha \) is set at 0.5. Accordingly, we can compute the incremental typicality (add decisions) and the relative incremental typicality (drop decisions):

<table>
<thead>
<tr>
<th>( C(I) )</th>
<th>( \exp (\ell(I)) )</th>
<th>Incremental Typicality</th>
<th>Max - Incremental Typicality</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \ell(1) = 6 )</td>
<td>403.43</td>
<td>( \ln (1 + 2.7183) = 1.3133 )</td>
<td>0.5433</td>
</tr>
<tr>
<td>( \ell(2) = 2 )</td>
<td>7.39</td>
<td>( \ln (1 + 0.0134) = 0.0133 )</td>
<td>-</td>
</tr>
<tr>
<td>( \ell(3) = 8 )</td>
<td>2980.95</td>
<td>( \ln (1 + 5.4018) = 1.8566 )</td>
<td>-</td>
</tr>
<tr>
<td>( \ell(4) = 5 )</td>
<td>148.41</td>
<td>( \ln (1 + 0.3679) = 0.3133 )</td>
<td>1.5433</td>
</tr>
<tr>
<td>( \ell(5) = 5 )</td>
<td>0</td>
<td>( \ln (1 + 0.2689) = 0.2382 )</td>
<td>-</td>
</tr>
</tbody>
</table>

Hence, the derived sequence of consideration is: 3 (add), 4 (drop), 1 (drop), 1 (drop), 5 (add), 2 (add).

Outcome: 3 cannot be added (incremental typicality < decision cost) \( 1.8566 < 2.3760 \)
4 dropped (relative incremental typicality < decision cost) \( 1.5433 < 6.5798 \)

With the new reference set configuration, we have

<table>
<thead>
<tr>
<th>( C(I) )</th>
<th>Incremental Typicality</th>
<th>Max - Incremental Typicality</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \ell(1) = 6 )</td>
<td>1</td>
<td>( \ln (1 + 403.43) = 6.0025 )</td>
</tr>
<tr>
<td>( \ell(2) = 2 )</td>
<td>0</td>
<td>( \ln (1 + 0.0183) = 0.0182 )</td>
</tr>
<tr>
<td>( \ell(3) = 8 )</td>
<td>0</td>
<td>( \ln (1 + 7.3890) = 2.1269 )</td>
</tr>
<tr>
<td>( \ell(4) = 5 )</td>
<td>0</td>
<td>( \ln (1 + 0.3679) = 0.3133 )</td>
</tr>
<tr>
<td>( \ell(5) = 5 )</td>
<td>0</td>
<td>( \ln (1 + 0.3679) = 0.3133 )</td>
</tr>
</tbody>
</table>

The sequence of consideration now is 3 (add), 4 (add), 5 (add), 2 (add), 1 (drop).

Outcome: 3 added (incremental typicality = 2.1269 > decision cost = 1.6450).

42
For the new reference set configuration, we obtain

<table>
<thead>
<tr>
<th>C (I)</th>
<th>Incremental Typicality</th>
<th>Max-Incremental Typicality</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t^f(1) = 6$</td>
<td>1</td>
<td>$\ln (1 + 0.1353) = 0.1269$</td>
</tr>
<tr>
<td>$t^f(2) = 2$</td>
<td>0</td>
<td>$\ln (1 + 0.0022) = 0.0022$</td>
</tr>
<tr>
<td>$t^f(3) = 8$</td>
<td>1</td>
<td>$\ln (1 + 7.3890) = 2.1269$</td>
</tr>
<tr>
<td>$t^f(4) = 5$</td>
<td>0</td>
<td>$\ln (1 + 0.0439) = 0.0429$</td>
</tr>
<tr>
<td>$t^f(5) = 5$</td>
<td>0</td>
<td>$\ln (1 + 0.0439) = 0.0429$</td>
</tr>
</tbody>
</table>

The derived sequence of consideration becomes: 1 (drop), 4 (add), 5 (add), 2 (add), 3 (drop).

**Outcome:** 1 dropped (incremental typicality = 0.1269 < decision cost = 0.2500).

No brand can be added or dropped subsequently, and the final composition of reference set C is: 0 0 1 0 0, and hence = {3}.

43
Appendix B: Likelihood of Being in the Choice Set \( P_i (j \in M^l) \)

The effect of reference set size, variance and mean of the typicality judgment on likelihood of being included in the choice set can be assessed as follows. By definition,

\[
b_i^l = \ln \left[ \sum_{j \in C} \exp \left( t_{ij}^l \right) \right]
\]

Using a Taylor expansion, \( b_i^l \) is approximately equal to

\[
b_i^l \approx \ln \left[ n + n t_i^l + \frac{1}{2} \sum_{j \in C} \left( t_{ij}^l \right)^2 \right] \tag{B-1}
\]

given \( n t_i^l = \sum_{j \in C} t_{ij}^l \) where \( n \) denotes the number of brands in reference set \( C \).

Furthermore,

\[
\sum_{j \in C} \left( t_{ij}^l \right)^2 = n s_i^2 + n t_i^l
\]

with \( s_i^2 = \sum_{j \in C} \left( t_{ij}^l - t_i^l \right)^2 / n \)

Moreover, expression (B-1) becomes

\[
b_i^l \approx \ln \left[ 1 + t_i^l + \frac{1}{2} t_i^l + \frac{1}{2} S_i^2 \right] n
\]

Taking partial derivatives, we obtain

\[
\frac{\partial b_i^l}{\partial n} = \frac{1}{n}
\]
which is always larger than zero.

Furthermore,

\[
\frac{\partial b^f_i}{\partial t^f_i} = \frac{1 + t^f_i}{1 + t^f_i + (1/2) t^f_i + (1/2) s^2_i}
\]

and

\[
\frac{\partial b^f_i}{\partial s^2_i} = \frac{1}{2(1 + t^f_i + (1/2) t^f_i + (1/2) s^2_i)}
\]

which are both positive when \( t^f_i \) is positive.

In sum, all three derivatives are generally positive which together with \( \frac{\partial P_i(j \in M^f)}{\partial b^f_i}(0) \) implies that \( P_i(j \in M^f) \) decreases as the number of prototypical alternatives, their mean typicality, and the variance in their typicality increases.

The effect of uncertainty in the typicality judgments can be assessed by adding an uncertainty parameter to the Type-1 Extreme Value distribution as follows

\[
p(\tau^f_{ij}) = \Theta_i^{-1}\exp\left[-\frac{\tau^f_{ij}}{\Theta_i} - \exp\left(\frac{\tau^f_{ij}}{\Theta_i}\right)\right]
\]

where the variance of the random component \( \tau^f_{ij} \) is proportional to parameter \( \Theta_i \). Hence, larger values of \( \Theta_i \) would indicate increased uncertainty in typicality. It is straightforward to show that

\[
\frac{\partial P_i(j \in M^f)}{\partial \Theta_i}(0)
\]

which implies that the likelihood increases with increased uncertainty.
Please forward your requests for working papers to the following address:

Executive Officer
Department of Marketing
School of Business and Management
Hong Kong University of Science & Technology
Clear Water Bay
Kowloon, Hong Kong

www: http://www bm ust h k/~mark/
Enquiries: Tel (852) 2358 7700 Fax (852) 2358 2429