ACTIVE VERSUS PASSIVE LOYALTY: A STRUCTURAL MODEL OF CONSIDERATION SET FORMATION

Nitin Mehta
Rotman School of Management
University of Toronto

Surendra Rajiv
Graduate School of Business
University of Chicago

Kannan Srinivasan
Graduate School of Industrial Administration
Carnegie Mellon University

April 1999
Revised: June 2001

The authors are listed in alphabetical order and contributed equally to the paper. We thank the Editor, the Area Editor and two anonymous reviewers for their constructive comments and suggestions. We also benefited from insightful comments from the seminar participants at Columbia University, NYU, Stanford University, and Universities of Chicago and Toronto. This work was partly supported by the Beatrice Companies Faculty Research Fund (at the GSB, University of Chicago) to the second author. The usual disclaimer applies.

ABSTRACT

We offer an econometric framework that models a consumer's brand choice decision as a two-stage process: consideration set formation followed by brand selection from the brands in the consideration set. The proposed structural model of consideration set is motivated by the fact that consumers have limited information-acquisition ability. In the context of frequently purchased products (FPPs), since these product categories are characterized by frequent price promotion of varying depths of discount, a consumer faces significant uncertainty about the net utilities associated with the different brands. Thus, while a consumer might know the range of potential prices, she is unaware of the actual posted price of a brand on any given purchase occasion unless she engages in pre-evaluation price search. Since information acquisition is costly, she needs to first decide how many and which brands to search the posted prices of. In the proposed framework, the process of pre-evaluation price search is conceptualized as "brand consideration" and the set of sampled brands is referred to as the consumer's "consideration set".

A distinctive feature of the proposed specification is that it allows us to distinguish between two sources of state dependence – viz., passive and active brand loyalty. In this conceptualization, *passive brand loyalty* refers to state dependence arising out of "consumer lock-in" as a result of high search cost. Thus, a passively loyal consumer repeatedly buys the same brand over successive purchase occasions because her cost of searching the posted prices of other brands is very high i.e. her optimal consideration set does not include any other brand. In contrast, *active brand loyalty* refers to state dependence arising out of a high intrinsic preference for the selected brand. Thus, while an actively loyal consumer considers more than the selected brand on any purchase occasion (because her search costs are low), she nonetheless buys the same brand because of her high intrinsic preference for the brand. Our key theoretical results are as follows: (i) relative to low price sensitive consumers, high price sensitive consumers have larger consideration sets; (ii) the intensity of consumer search is higher in product categories characterized by greater price variability. Thus, more frequent price promotions with deep discounts lead to large consideration sets; and, (iii) a consumer does not stay in the state of inertia for long. In the case of a passively loyal consumer, we will observe phases of inertia followed by a brand switching that in turn is followed by another spell of inertia.

We use scanner panel data for liquid detergents to empirically validate the model. The key empirical results are as follows: (i) there are significant search costs that consumers incur in discovering the actual posted prices of the brands at the store. This implies that consumers do not consider (i.e. search the posted prices of) all the brands on a shopping trip; (ii) in-store display activities and feature ads reduce consumer search costs for a brand thereby significantly increasing the probability of the brand being considered. Feature advertising reduces the search costs more than in-store displays; (iii) prior consumption influence quality perception of a brand for both liquid detergents and ketchup categories. However, consumption experience yield only limited additional quality information; and, (iv) estimates of price sensitivity critically depend on whether the specification explicitly models consideration stage. In particular, a model that assumes that consumers search all the brands on all purchase occasions seriously underestimates the impact of price on brand choice decision. We also conduct a limited cross-category analysis using ketchup data set and find several interesting differences in consumer price search behavior across the two product categories analyzed.

KEY WORDS: *Consumer Inertia; Consideration Set; Brand Loyalty; Consumer Learning; Bayesian Updating; Structural Model.*

1. INTRODUCTION

Since the seminal work of Howard and Sheth (1969), the concept of "consideration set" (also referred to as "evoked set" or "choice set" in the literature) has gained considerable acceptance. Consideration set refers to the set of brands (a subset of all the brands in the product category) over which a consumer makes an explicit utility comparison or cost-benefit trade-off before she makes her brand choice decision. The notion of consideration set has been rationalized in the literature on the basis of either limited information-processing ability or limited information acquisition ability on the consumer's part (Manrai and Andrews, 1998). The basic idea behind the first rationale is that a consumer has limited information-processing abilities and hence she cannot make explicit utility comparisons across all the brands (since that will entail her processing attribute information about all the brands). To reduce her processing task, the consumer limits this comparison to a subset of brands that is termed as her consideration set (e.g., Nedungadi, 1990). Note that this explanation implicitly assumes that all the relevant attribute information about all the brands are stored in the consumer's mind; the only issue then is the cost of retrieving and processing this information. In contrast, the limited-information-acquisition-ability rationale for consideration sets involves the notion of consumer uncertainty about the surplus associated with the different choice alternatives and, hence, the need for information-search (e.g., Ratchford, 1982; Feinberg and Huber, 1996). The idea here is that a consumer is uncertain about the attributes of the alternative brands and hence must actively gather information about these brands before making an explicit trade-off to arrive at brand choice. Since information acquisition is a costly and time-consuming process, the consumer approaches the decision-making task in a sequential manner that is described as follows:

- First, she decides the set of brands about which she will gather more detailed information so as to get a more precise estimate of the associated consumer surplus; and,
- Second, having obtained this information, she decides the utility-maximizing brand based on an explicit trade-off between these considered brands.

In sharp contrast to this conceptualization of consumer brand choice as a phased decision-making process, prior stochastic choice models, barring a few exceptions that we discuss in section 1.2, typically assume that consumers employ all the relevant information about all the brands in the product category on all the purchase occasions. Thus, in these models, consumers are assumed to be perfectly informed about the posted prices and other marketing-mix variables of all the brands; further, consumer brand choice decision is arrived at by explicit utility comparison across all the brands in the product category. In addition, since these models take a reduced-form approach, they do not explicate the mechanisms through which point-of-purchase marketing activities such as feature advertising and display activities affect consumer brand choice.

Our objective in this paper is to propose a structural model of consideration set formation, based on the limited-information-acquisition-ability motivation alluded to earlier. If consumers are uncertain about the net utilities associated the different brands, we need to address the following issues:

- What are the sources of this uncertainty about consumer surplus associated with the brands?
- What sources of uncertainty get resolved as a result of consumer information acquisition?
- What factors might influence the costs associated with information acquisition?
- What strategies and decision rules do consumers adopt while deciding how much and what type of information to acquire?

In the context of consumer durables (and other infrequently purchased high-ticket items), consumers can acquire information about the attributes of the different brands through explicit search, for example, through Consumer Reports, consulting experts etc. (Newman and Staelin, 1972). This, of course, assumes that the product category in question is a "search good" (Nelson, 1974). However, in the context of frequently purchased products (FPPs), consumers can learn about product quality through consumption experience since these products are more "experience or credence goods" (Kotler, 1996). Since the FPPs are frequently promoted, consumers do not have complete information about their price: while they may know the range of possible prices, they do not know the actual posted prices on a particular shopping trip. This leads to uncertainty in consumer surplus associated with the different brands. However, by engaging in explicit price search, consumers can ascertain the posted prices and thereby reduce uncertainty.

1.1 Brief Overview of the Econometric Specification and Main Empirical Findings

The key components of the proposed specification are as follows. Consumers are uncertain about the prevailing prices of the brands, though they are aware of the price distributions. Consumer rationality implies that consumers will engage in price search to reduce this uncertainty. However, since price search is costly, consumers make a trade-off between higher expected utility from a more extensive price search and the cost of search. *In the proposed specification, we conceptualize the costs associated with price search as the "consideration costs" that lead a consumer to sample only a subset of brands in the product category. This subset of sampled brands is conceptualized as the consumer's "consideration set"*. The consideration set formation and the final choices are modeled probabilistically. The optimal consideration set and search costs are endogenously estimated from scanner panel data set. Also, consumers are uncertain about quality of the brands in the product category and they update their beliefs through consumption experiences in a Bayesian fashion. This component of the model draws on the approach proposed by Erdem and Keane (1996).

A distinctive feature of the proposed specification is that it allows us to distinguish between two sources of "state dependence"[1] – viz., passive and active brand loyalty. In this conceptualization, *passive brand loyalty* refers to state dependence arising out of "consumer lock-in" as a result of high search costs. Thus, a passively loyal consumer repeatedly buys the same brand over successive purchase occasions because her costs of searching the posted prices of other brands is very high i.e. her optimal consideration set does not include any other brand. Thus, passive brand loyalty is akin to the notion of consumer inertia identified in the prior literature (Jeuland, 1979; Wernerfelt, 1991). In contrast, *active brand loyalty* refers to state dependence arising out of a high intrinsic preference for the selected brand. Thus, while an actively loyal consumer considers more than the selected brand on any purchase occasion due to her low search costs, she nonetheless buys the same brand because of her high intrinsic preference for the brand.

Our structural model yields the following insights about the size of the optimal consideration set and consumers' inertial behavior in frequently purchased product categories:

(a) Relative to low price sensitive consumers, high price sensitive consumers have larger consideration sets.

(b) The intensity of consumer search is higher in product categories characterized by greater price variability. Thus, more frequent price promotions with deep discounts lead to large consideration sets.

(c) A consumer does not stay in the state of inertia for long. In the case of a passively loyal consumer, we will observe phases of inertia followed by a brand switching that in turn is followed by another spell of inertia.

We calibrate the proposed model using ERIM data set for liquid detergent product category. To analyze cross-category variations in consumer price search behavior, we also calibrate the model using ketchup data set. Our empirical analysis yields the following key results:

(i) There are significant search costs that consumers incur in discovering the actual posted prices of the brands at the store. This implies that consumers do not consider (i.e. search the posted prices of) all the brands on a shopping trip. For liquid detergents, we find that the average consideration set size is 2.31. For ketchup category, we find that consumers undertake more extensive price search with the average consideration set size being 2.47.

(ii) In-store display activities and feature ads reduce consumer search costs for a brand thereby significantly increasing the probability of the brand being considered. Further, we find that for liquid detergents, feature ads are more effective than displays in influencing consumer search behavior. The opposite holds true for ketchup product category.

1. State dependence refers to the observed pattern in many frequently purchased products (FPP) categories wherein a

(iii) Prior consumption influence quality perception of a brand for both liquid detergents and ketchup categories. However, consumption experience yield only limited additional quality information. This is consistent with our expectation since these categories are mature.

(iv) Estimates of price sensitivity critically depend on whether the specification explicitly models consideration stage. In particular, a model that assumes that consumers search all the brands on all purchase occasions seriously underestimates the impact of price on brand choice decision.

1.2 Related Literature and Research Contributions

Some recent studies have attempted to incorporate the notion of consideration or choice set in the context of stochastic choice models. In a pioneering research effort, Hauser and Wernerfelt (1990) propose a theoretical evaluation cost model of consideration set. The basic idea here is that consumers face uncertainty about the utilities associated with the various brands and further, that by engaging in a costly pre-evaluation search consumers can resolve (or at least reduce) this uncertainty about brand utilities. Thus, the consumer first decides how many brands to search and then evaluates these brands based on the information gathered at the search stage to make a brand choice. Roberts and Lattin (1991) propose the first empirically estimable model of consideration set formation, based on the Hauser-Wernerfelt (1990) framework. The limitation of their econometric specification is that survey data, in addition to scanner panel data, is needed to calibrate the model. Andrews and Srinivasan (1995) extend the Roberts-Lattin (1991) specification to allow for probabilistic consideration sets that makes it possible to calibrate the model using scanner panel data alone. Bronnenberg and Vanhonacker (1996) extend the Andrews and Srinivasan (1995) framework by incorporating unobserved consumer heterogeneity using latent class approach. They also make an important observation that consideration set composition does not depend on actual posted prices but by the brands' price range membership. Siddarth, Bucklin and Morrison (1995) develop an alternative reduced-form specification of probabilistic consideration set that entails estimating individual-level choice sets using a Bayesian updating procedure in conjunction with the multinomial Logit model. Chiang, Chib and Narasimhan (1999) have recently proposed a model that accounts for consumer heterogeneity in consideration sets and in the parameters of the brand choice model.

The extant econometric models of consideration set formation, being reduced form specifications, do not explicate the driving force behind consideration. For example, Andrews and Srinivasan (1995) and Bronnenberg and Vanhonacker (1996) posit that all the brands whose "consideration utility" exceeds a certain threshold form part of the consumer's consideration set and further that a brand's "consideration utility" is a function of the marketing-mix variables. Thus, it is not immediately clear as to what consideration

---

consumer repeatedly buys the same brand over successive purchase occasions.

really means and why should “consideration utility” be different from “choice utility”. Similarly, the specification due to Chiang, Chib and Narasimhan (1999) takes a purely combinatorial approach to consideration sets and is not motivated by any behavioral assumption. Further, while this framework allows for consumer heterogeneity in considerations sets, for an individual consumer consideration set remains stable throughout her consumption history. However, there is sufficient experimental (e.g., Mitra, 1995) and empirical evidence (e.g., Allenby and Ginter, 1995) to suggest that consideration sets are not stable and are influenced by situational factors.

Another stream of literature has attempted to incorporate, in a reduced-form approach, the notion of limited consumer search. Bucklin and Lattin (1991) propose a two-state model of consumer brand choice that explicitly models two types of shopping behavior viz., planned and unplanned purchases. They posit that while for unplanned purchases, consumers are influenced by marketing-mix variables, for planned purchases, consumers do not process in-store marketing-mix information and hence are not influenced by point-of-purchases interventions such as price promotion. Recently, Murthi and Srinivasan (1999) propose a complete- and limited-information model of brand choice and observe that there is considerable evidence that search cost substantially reduces the extent of evaluation of product category information. Note that in both these models, consumers either engage in searching marketing mix information about all the brands in the product category or they do not search information on any brand (though Murthi and Srinivasan (1999) allow for consumers to process information on only a limited set of marketing-mix variables).

We view our main methodological contribution as being the integration of these two research streams. The proposed model of consideration sets is developed taking the limited information acquisition ability perspective and the notion of limited price search is central to our conceptualization of consideration set formation. Derived from the primitives of consumer utility maximization, our formulation also provides insights into variation in consumer search behavior across product categories in terms of consumer characteristics (e.g., price sensitivity) and firms’ behavior (e.g., price promotion policy).

The rest of the paper is arranged as follows. In Section 2, we develop the model formulation and discuss the estimation method. Section 3 describes the data set used for calibrating the proposed model, reports the parameter estimates and discusses the managerial implications of our findings. Section 4 concludes.

2. MODEL DEVELOPMENT

In this section, we layout the model specification and discuss the main assumptions underlying the proposed formulation. Before developing the econometric specification, in section 2.1 we provide a conceptual description of the model. Section 2.2 lays out the mathematical details underlying the proposed specification.[2] Section 2.3 discusses our approach to control for unobserved consumer heterogeneity. Section 2.4 discusses the necessity for explicitly modeling consumer quality-learning in order to derive an identifiable statistical model of consideration set formation. Section 2.5 distinguishes between preference-based or "active" loyalty and consumer inertia or "passive" loyalty. Finally, section 2.6 provides insights into some of the key determinants of the extent of consumer price search (and consequently the size of the optimal consideration set). Specifically, we provide the comparative statics results for the size of the consideration set with respect to (i) consumer price sensitivity; and, (ii) the variability of prices.

2.1 Conceptual Description

The basic ingredients of the proposed model of consideration set formation are as follows:

(i) At the beginning of her consumption history, a consumer is uncertain about the intrinsic quality (and hence the associated utility) of the brands in the product category; this initial quality uncertainty is captured by hypothesizing that the consumer has prior beliefs about quality of the brands in the product category.

(ii) The consumer learns about the intrinsic quality of the brands in the product category through consumption experiences. Further, consumption experience does not fully reveal the intrinsic brand qualities. Said differently, consumption experience only provides a *noisy signal* about the brand's intrinsic quality. Section 2.2.1 describes the Bayesian updating mechanism underlying the evolution of quality perceptions over a consumer's purchase history. Note that while brand qualities are uncertain and evolve over time, we assume that consumers cannot learn additionally about quality of the various brands while at the store through any conscious search/information acquisition effort.

(iii) Since the brands in the product category are occasionally on "sale" and further since the depth of discount offered during price promotions vary across sale events, a consumer is uncertain about the posted price of the brands on any purchase occasion. However, we assume that through prior

2. Analytical details are more elaborately discussed in the Technical Appendix that is available from the authors on request.

observation the consumer is aware of the distribution of prices for the various brands that are assumed to be stationary.[3]

(iv) Before making her brand choice decision, the consumer can ascertain the posted prices of all the brands in the category or a subset thereof. However, this price search is costly to the consumer because it entails investment in time and effort. This implies that while deciding the number of brands to search on a given purchase occasion, a utility maximizing consumer makes a trade-off between the potential gains from searching an incremental brand and the associated search costs. Thus, in the proposed formulation, *"consideration" is conceptualized as the process of price search and "consideration set" refers to the optimal subset of brands whose prices the consumer selects to search*. Section 2.2.2 provides the analytical details of the consumer trade-off and the process of optimal consideration set formation.

(v) Given her consideration set, the consumer observes the posted prices of the included brands. The consumer then selects the brand that offers the highest expected consumer surplus (Note that while the price uncertainty gets resolved as a result of price search, the consumer continues to remain uncertain about the intrinsic brand qualities at the brand choice stage). Section 2.2.3 derives the brand choice probabilities based on the random utility maximization paradigm (McFadden, 1981).

2.2 Model Formulation

2.2.1 Utility Specification & Evolution of Quality Perceptions

Consider a product category with $j = 1, \ldots, J$ brands. We assume that consumer i's (indirect) utility or surplus from brand j on purchase occasion t can be approximated as a linear function of brand j's perceived quality, q_{ijt}, and price, p_{ijt}, as follows:

$$U_{ijt} = \theta q_{ijt} - p_{ijt}\,. \tag{1}$$

The parameter θ denotes the consumer's intensity of preference for quality (alternatively, her marginal willingness-to-pay for quality). The above specification for consumer surplus function has been traditionally employed in the vertical differentiation literature (e.g., Moorthy, 1988; Bagwell and Riordan, 1991). Further, this is similar to the specification used by Erdem and Keane (1996) in their quality-learning model.[4]

---

3. Clearly, this is a simplification for reasons of analytical tractability. One could potentially allow consumers to update their beliefs about the distribution of brands in their consideration sets.

4. Specifically, Erdem and Keane use the specification: $U_{ijt} = q_{ijt} - \gamma p_{ijt}$. Both of these specifications essentially capture a consumer's quality-price trade-off. In our model, it is captured through the parameter θ while in Erdem and Keane (1996) formulation, γ – being the price sensitivity parameter – captures this trade-off. Both these specifications of indirect utilities are derived assuming no income effect and weak separability of preferences. It is

We further assume that consumers are uncertain about the intrinsic qualities of the competing brands. This is modeled by positing that consumer *i* on any purchase occasion *t*, instead of being aware of the "true" quality of brand *j*, holds only *subjective quality beliefs* which is captured by her priors $f(q_{ijt})$. In particular, we assume a normal distribution for consumer prior beliefs i.e.,

$$q_{ij,t} \mid H_i(t) \sim N\left(\omega_{ij,t}, \sigma^2_{\omega_{ijt}}\right), \quad \forall i, \forall j. \tag{2}$$

Note that $\omega_{ij,t}$ denotes consumer *i*'s current estimate of the expected quality of brand *j* given her consumption history, $H_i(t)$. Further, $\sigma^2_{\omega_{ijt}}$ denotes the extent of consumer uncertainty about brand *j*'s quality on purchase occasion *t*.

We assume that prior to any consumption experience in the product category (i.e., at purchase occasion $t = 0$), the consumer holds identical quality beliefs about all the brands in the product category. Further, this belief is assumed to be identical across all consumers. Thus, we assume that initial quality beliefs are given by

$$q_{ij,t} \mid H_i(t=0) \sim N\left(\omega_0, \sigma^2_{\omega,0}\right), \quad \forall i, \forall j.^{5} \tag{3}$$

We assume that a consumer learns about the intrinsic quality of a brand through consumption experiences. We further assume that consumption experience provides only partial information about the brand's intrinsic quality. In other words, consumption experience is only a *noisy signal* of a brand's intrinsic quality. The rationale for this assumption is as follows. First, the noise in consumption experience could arise because of "inherent product variability" (Roberts and Urban, 1988; Moorthy and Srinivasan, 1995) i.e., the delivered quality of a brand fluctuates about its intrinsic quality. Second, consumers may misjudge the quality realization for the brand on a particular consumption occasion. The net impact of this assumption is that while a consumer's uncertainty about a brand's quality reduces over her consumption history, the "true" quality of a brand does not get fully revealed (Erdem and Keane, 1996).[6]

Let $q_{ij,t-1} \sim N\left(\omega_{ij,t-1}, \sigma^2_{\omega_{ij,t-1}}\right)$ be the *i*th consumer's perceived quality of brand *j* on purchase occasion *t*-1. Further, let $\lambda_{ij,t-1}$ denote the (noisy) quality cue associated with consumption experience on *t*-1 occasion with

---

the same set of assumptions needed for the Logit and the Probit models to be consistent with random utility maximization (McFadden, 1981).

5. We make this assumption for expositional simplicity. In our empirical analysis, we use an initialization sample to estimate consumers' initial quality beliefs and use these for the estimation sample.

6. In this sense, the product category is more a credence good rather than an experience good (Horstmann and MacDonald, 1994; Kotler, 1996).

$$\lambda_{ij,t-1} = q_j + \eta_{ij,t-1}. \tag{4}$$

In equation (4), q_j denotes the "true" quality of brand *j* while the random variable $\eta_{ij,t-1}$ denotes the noise associated with consumption experience with $\eta_{ij,t-1} \sim N\left(0, \sigma_\eta^2\right)$.[7] Thus, by consuming brand *j* in period *t*-1 the consumer gets to learn more about the "true" quality of the brand. However, since the signal is noisy, her expectation about brand *j*'s quality, $\omega_{ij,t-1}$, does not get updated to q_j. Note that σ_η^2 is a measure of the informativeness of consumption experience: σ_η^2 = 0 corresponds to the product category being a classic experience goods so that the consumer gets to learn the "true" quality after a single consumption experience (Nelson, 1974, Milgrom and Roberts, 1986). For analytical tractability, we assume the variance of the noise term, σ_η^2, to be the same across all the brands and over time.

We assume that the consumer uses the quality cue from consumption experience, $\eta_{ij,t-1}$, to update her prior beliefs about brand *j*'s quality, $q_{ij,t-1}$, using Bayes rule. Let $d_{ij,t-1}$ be an indicator variable such that $d_{ij,t-1}$ = 1 if consumer *i* purchases brand *j* on purchase occasion *t*-1. Using the fact that normal density is self-conjugate (DeGroot, 1970), it can be shown that the mean of consumer *i*'s quality perception of brand *j*, given her consumption history, $\mathrm{H}_i(t)$, is given by[8]

$$\omega_{ij,t} \quad = \quad \omega_{ij,t-1} \quad + \quad N\left(0,\ \frac{d_{ij,t-1}\left(1+\alpha_{ij,t-1}{}^{-1}\right)}{\left(\alpha_{ij,t-1}+d_{ij,t-1}\right)^2}\right). \tag{5}$$

In equation (5), we use $\alpha_{ij,t-1} = \sigma_\eta^2 \big/ \sigma_{\omega_{ij,t-1}}^2$.[9] We can interpret $\alpha_{ij,t}$ as the inverse of signal to noise ratio for brand *j*'s consumption experience. It measures the extent to which the uncertainty of a consumer exposed to successive *t*-1 consumption experiences differs from the "inherent product variability". In particular, a high value of $\alpha_{i,t}$ would suggest that additional consumption experiences are not very effective in further influencing her quality perception of brand *j*.

---

7. The assumption that the noise associated with consumption experience is normal is made to make use of the fact that normal density is self-conjugate.

8. The analytical details are given in the Technical Appendix. In technical terms, $\{\omega_{ij,s}\}_{s=1}^{T}$ is a Martingale sequence with constant mean and $E\left[\omega_{ij,t} \mid \mathrm{H}_i(t)\right] = \omega_{ij,t-1}$.

9. Since we can only estimate $\alpha_{ij,t-1} = \sigma_\eta^2 \big/ \sigma_{\omega_{ij,t-1}}^2$, for identification purposes, we set the variance of consumption signal to be one i.e. σ_η^2 = 1.

Note that the second term in equation (5), subsequently denoted compactly as $N_{ij,t-1}$, is a normal random variable that can take both positive and negative values. Thus, as in Erdem and Keane (1996), the proposed specification allows a very flexible quality learning mechanism in which individual consumption experiences can have both favorable and unfavorable impact on quality perception.[10] This is distinct from the reduced-form operationalizations of "purchase-event feedback" or brand loyalty (e.g., Guadagni and Little, 1983) that allows for only positive (or negative as in Seetharaman, Ainslie and Chintagunta, 1998) updating of quality perceptions over successive consumption experiences.

We assume the initial values of $\omega_{ij,t}$ and $\alpha_{ij,t}$ to be the same across all brands and across all consumers and without loss of generality set it to zero, i.e., $\omega_{ij,0} = 0$ and $\alpha_{ij,0} = \alpha_0$, $\forall i, \forall j$. Thus, from equation (5), consumer *i*'s mean quality perception of brand *j* given her purchase history $\mathrm{H}_i(t)$, is given by $\omega_{ij,t} = \sum_{l=0}^{t-1} N_{ij,l}$. It is important to note here that the mean quality perception of brand *j*, $\omega_{ij,t}$ – being a sum of the realized values of *t* random variables – is *not* a random variable from the consumer's perspective since she gets to observe these realized values of the consumption signals viz., $\lambda_{ij,t-1}$, $\forall t$. However, from the perspective of the researcher, $\omega_{ij,t}$ is a random variable that captures that part of the consumer's preference that is unobservable and time varying.[11] This is analogous to the "random error component" in reduced-form specifications such as Logit (e.g. Guadagni and Little, 1983) or Probit (e.g., Chintagunta, 1993) models. Another implication of equation (5) is that the random error component exhibits serial correlation i.e. "habit persistence" (Roy, Chintagunta and Haldar, 1996). Further, since the variance of $\omega_{ij,t}$ evolves over time, it is heteroskedastic as well.

In short, the proposed quality-evolution specification provides an intuitive interpretation of the "statistical error term" in the reduced-form choice models. It further raises questions about the standard assumptions of homoskedasticity and absence of serial correlation routinely made in the literature. Note

---

10. In fact, as we discuss in section 2.5, it is this fact that allows us to distinguish between active and passive loyalty.
11. The researcher has some information about the expected qualities $\omega_{ij,t}$. For example, the distribution function of $\omega_{ij,t}$ is known. In addition, the choices made by consumer *i* over her consumption history (i.e. from *t*=0 to *t*-1) is also known. This gives the researcher the information about the relative magnitudes of the quality signals that the consumer had received in every time period till time *t*-1. This in turn gives the researcher the information about relative magnitudes of $\omega_{ij,t}$ for all the *j* brands.

further that we do not need to incorporate an additional error term (e.g., as in Roberts and Lattin, 1991 or Erdem and Keane, 1996). In fact, such an error term will be unidentifiable in this specification.[12]

2.2.2 Determination of the Optimal Consideration Set

Consider the problem being faced by consumer *i* on purchase occasion *t*. Based on her consumption experience, $\mathrm{H}_i(t)$, her updated (posterior) beliefs about qualities of the competing brands are denoted by $q_{ij,t} \sim N\left(\omega_{ij,t}, \sigma^2_{\omega_{ij,t}}\right),\ \forall j$.

In addition to quality uncertainty, the consumer faces uncertainty about the posted prices of the various brands on that purchase occasion. The basic idea here is that since retailers routinely offer price promotions of varying depths of discount, the actual price prevailing on any shopping trip is a random variable. While the consumer does not know the actual posted price, we assume that she is aware of price distributions for the various brands denoted by $p_{ijt} \sim f_j\left(\bar{p}_j, \sigma^2_{p_j}\right),\ \forall j$, where $\bar{p}_j$ is the mean or expected price and $\sigma^2_{p_j}$ is the variability in brand *j*'s price. Note that $\sigma^2_{p_j}$ increases with the frequency of price promotions and the depths of discount offered when brand *j* is on "sale". To ensure that the net surplus or (indirect) utility function $U_{ij,t}$ has a Type 1 extreme value (EV) distribution, we assume that the price of brand *j* has a Type 1 EV distribution. We further assume that the price distributions are stationary and do not change with time. These assumptions are needed to ensure a closed-form expression for the expected maximum utilities given in equation (6).

Given the two-fold uncertainty on both the intrinsic quality as well as the prevailing price, consumer surplus associated with brand *j*, $U_{ijt} = \theta q_{ijt} - p_{ijt}$, is a random variable with mean $\theta\omega_{ij,t} - \overline{p}_{ij,t}$. In particular, the expected (indirect) utility associated with brand *j*, prior to the consumer discovering the actual posted price, $p_{ij,t}$, is a random EV variable with mean $\theta\omega_{ij,t} - \overline{p}_{ij,t}$ and variance $\sigma^2_{uj} = \sigma^2_{pj}$ i.e.,

$$U_{ij,t} \quad \sim \quad EV\left(\frac{\pi}{\sqrt{6}\sigma_{uj}},\ \frac{\sqrt{6}\sigma_u e_c}{\pi} - \theta\omega_{ij,t} + \overline{p}_j\right), \tag{6}$$

where e_c is the Euler's constant.[13]

---

12. Specifically, adding an i.i.d. error term with zero mean and finite variance would make the model unidentifiable. While we can introduce the additional error term assuming a known variance, as in Erdem and Keane (1996), it leads to estimation complexity without yielding any additional insights from the model.
13. We provide the analytical details in the Technical Appendix.

While the consumer cannot further reduce uncertainty surrounding brand *j*'s quality during the shopping trip, she recognizes that she can ascertain the actual posted prices (and thereby eliminate price uncertainty) by engaging in search. Now, if price search were costless, the consumer's optimal decision would be to search the posted prices of all the brands in the product category. Then, having ascertained the actual posted prices, she would choose the brand that yields the highest expected surplus (Note that even after the consumer has ascertained the posted prices, she still faces uncertainty about the net utility derivable from a brand due to quality uncertainty). However, if searching for the posted prices involves costs, then searching for the posted prices of all the brands in the product category may not be an optimal decision. In the proposed specification, *the process of price search is conceptualized as "brand consideration" so that the notion of "consideration set" is synonymous with the set of brands whose prices the consumer actually samples on that purchase occasion*.

This raises the question: how does the consumer decides how many and which brands to consider on any given occasion? In this paper, we assume that the consumer's optimal consideration decision is based on expected utility maximization and arrived at by making explicit trade-off between the costs and benefits from price search. Let $C_{ij,t}$ be the search cost that consumer *i* has to incur if she were to search the posted price of brand *j* on purchase occasion *t*, given her consumption history, $\mathrm{H}_i(t)$. The decision on the optimal size and composition of her consideration set would involve a trade-off between the benefit from including an additional brand *j'* and the additional cost $C_{ij',t}$ that she has to incur. Thus, the consumer selects those brands in her consideration set that yields the highest net expected surplus (net of the total search cost).

In order to derive the probability that any given set of brands, denoted by $\{k\}$, constitutes the consumer's optimal consideration set, we need to compute the expected (indirect) utility associated with this set and the total cost of searching for the posted prices of the brands in the set. We need to do this computation for all possible sets of brand $\{j\}$. Note that the expected (indirect) utility associated with any set $\{j\}$ is the same as the expected (indirect) utility of the "best" brand in that set. Clearly, since quality perceptions evolve over the consumer's purchase history, the evaluation also changes over successive purchase occasions.

Now, the maximum of a set of EV random variables is also an EV random variable (Johnson and Kotz, 1974). In order to get a closed-form expression for the expected utility of the "best" brand in the consideration set, we assume that the variance of price distribution $\sigma_{p_j}^2$ is same across all the brands in the product category. The net impact of this assumption is that the variance of the (indirect) utility function, $U_{ij,t}$, in equation (6) is the same across all the brands and remains constant over a consumer's purchase history i.e. $\sigma_{ujt}^2 = \sigma_u^2 = \sigma_p^2, \ \forall i, \forall j, \forall t$.[14]

Consider hypothetically the consumer selecting brands 1 to *N* with $N \le J$ in her consideration set. It can be shown that the expected benefit (net of search cost) of including brands 1 to *N* in the consideration set $\{j\}$ is given by[15]

$$EB_{i,t}^{\{j\}} \quad = \quad \frac{\sqrt{6}\sigma_u}{\pi} \ \ln\left(\sum_{l=1}^{N} \exp\left(\frac{\pi}{\sqrt{6}\sigma_u}\left(\theta\omega_{il,t} - \overline{p}_j\right)\right)\right) \ - \ \sum_{l=1}^{N} C_{il,t} \,. \tag{7}$$

Note that the expected benefit in equation (7) is concave in *N*, the number of brands in the consideration set. Thus, the optimal size and composition of the consideration set $\{k\}$ is given by

$$\{k\} \quad = \quad \underset{\{j\}}{\operatorname{argmax}} \ \frac{\sqrt{6}\sigma_u}{\pi} \ \ln\left(\sum_{l \in \{j\}} \exp\left(\frac{\pi}{\sqrt{6}\sigma_u}\left(\theta\omega_{il,t} - \overline{p}_l\right)\right)\right) \ - \ \sum_{l \in \{j\}} C_{il,t} \,. \tag{8}$$

In equation (8), since the expected qualities $\omega_{il,t}$ is known to the consumer, she can determine which of the sets $\{j\}$ yield the highest expected benefit. However, from the researcher's perspective $\omega_{il,t}$ are random variables (since the researcher does not observe the realized values of the quality cues). As such, given the consumer's consumption history at time *t*, $\mathrm{H}_i(t)$, we can only write the probability that the set $\{k\}$ is the optimal consideration set:

$$\begin{aligned} \Pr\left[\{k\} = \text{Consideration Set}\right] \quad = \quad & \Pr\left[\frac{\sqrt{6}\sigma_u}{\pi} \ \ln\left(\sum_{l \in \{k\}} \exp\left(\frac{\pi}{\sqrt{6}\sigma_u}\left(\theta\omega_{il,t} - \overline{p}_l\right)\right)\right) \ - \ \sum_{l \in \{k\}} C_{il,t} \ge \right. \\ & \left. \frac{\sqrt{6}\sigma_u}{\pi} \ \ln\left(\sum_{l \in \{j\}} \exp\left(\frac{\pi}{\sqrt{6}\sigma_u}\left(\theta\omega_{il,t} - \overline{p}_l\right)\right)\right) \ - \ \sum_{l \in \{j\}} C_{il,t}\right], \qquad \forall \{j\} \,. \end{aligned} \tag{9}$$

2.2.3 Consumer Brand Choice Decision

Once the consumer has selected her optimal consideration set, $\{k\}$, she gets to observe the posted prices of all the brands included in her consideration set. That is, the price uncertainty gets completely resolved for all the brands under consideration, prior to the consumer making a brand choice. However, she still

---

14. Even if the variances of the indirect utilities associated with the brands were different, we would still be able to estimate our model. However, we won't have a closed-form expression for consideration probability (equation 9) that greatly simplifies the estimation. In the context of our empirical analysis, for the liquid detergent category, we find the variance of the prices to be almost the same across the 4 brands (see Table 1A). As a result, we have used equation (7) to compute the expected benefit for liquid detergents. In contrast, for the ketchup data set, the price variances are not similar (see Table 1B) and hence we have numerically computed the expected benefit to compute consideration probabilities.
15. Analytical details are given in the Technical Appendix.

remains uncertain about the intrinsic qualities of these brands. She then selects the brand in her consideration set that gives her the highest expected (indirect) utility. As before, since $\omega_{il,t}$'s are random variables from the researcher's perspective, we can only make probabilistic statement. In particular, since $\omega_{il,t}$'s are normal random variables (c.f. equation 5), the probability that the consumer selects brand *j*, given that $j \in \{k\}$, is given by a Probit probability:

$$\Pr\left(\delta_{ij,t}=1 \mid j\in\{k\}\right) = \Pr\left(\theta\omega_{ij,t}-p_{ij,t} \geq \theta\omega_{ih,t}-p_{ih,t}, \quad \forall\, h\in\{k\} \mid H_i(t)\right)$$

$$= \int_{-\infty}^{p^1_{ij,t}/\theta} \cdots \int_{-\infty}^{p^N_{ij,t}/\theta} \phi\left(\omega^1_{ij,t},\cdots,\omega^{j-1}_{ij,t},\omega^{j+1}_{ij,t},\cdots,\omega^N_{ij,t}\right) \prod_{\substack{h\in\{k\}\\ h\neq j}} d\omega^h_{ij,t} \qquad (10)$$

In equation (10), the indicator variable $\delta_{ij,t}$ = 1 if consumer *i* selects brand *j* on purchase occasion *t* and $\delta_{ij,t}$ = 0 otherwise. Further, we define the variables $\omega^h_{ij,t}$ and $p^h_{ij,t}$ as follows: $\omega^h_{ij,t}=\omega_{ih,t}-\omega_{ij,t}$ and $p^h_{ij,t}= p_{ih,t}-p_{ij,t}$.

Thus, the unconditional probability that brand *j* is selected is obtained by deriving brand choice probability given all possible consideration sets containing the candidate brand and then unconditioning over these possible consideration sets:

$$\Pr\left(\delta_{ij,t}=1\right) = \sum_{\{k\}} \Pr\left(\delta_{ij,t}=1 \mid j\in\{k\}\right)\times \Pr\left[\{k\}=\text{Consideration Set}\right] \qquad (11)$$

2.3 Controlling for Unobserved Consumer Heterogeneity

In the proposed formulation, since different consumers go through idiosyncratic consumption experiences, the quality evaluations are *not* the same across all consumers (as well as not the same across all the brands for the same consumer). In effect, this heterogeneity in quality beliefs corresponds to a richer specification of intercept heterogeneity than that in the literature (e.g., Chintagunta, Jain and Vilcassim, 1993; Gonul and Srinivasan, 1995) as it allows for:

- Unobserved heterogeneity in intrinsic brand preference across consumers; and,
- Unobserved heterogeneity in intrinsic brand preference across purchase occasions for the same consumer.

We are able to accomplish such finer variance decomposition because ω_{ijt} needs to satisfy certain conditions over the consumption history H_i that are discussed in detail in the Technical Appendix.

To account for heterogeneity in consumer's intensity of preference (marginal willingness-to-pay) for quality, we assume θ to be gamma distributed across the consumer population with mean $\bar{\theta}$ and variance σ_θ^2. This ensures that the consumer's marginal willingness-to-pay for quality is positive, consistent with the vertical differentiation literature.[16]

2.4 Quality Uncertainty & Model Identification

As noted earlier, the randomness in $\omega_{ij,t}$ (from the researcher's perspective) is the source of stochasticity in our formulation and is akin to the random utility component in the standard discrete choice models. If we were to use the "statistical error term" (e.g., as in Lattin and Roberts, 1991) instead of quality learning to build our statistical model, then we will not be able to identify and estimate all the parameters at the consideration stage in our framework. Specifically, we would not be able to separately identify the base-line search cost C_0 (the component of search cost that is not time varying) and the true qualities of the brands q_j. The reason is as follows: If there were no quality learning, the consumer would know the true quality of all the brands. The indirect utility of brand *j* in that case would be $U_{ijt} = \theta q_j - p_{ijt} + \varepsilon_{ijt}$. The probability for considering any set $\{k\}$ would be

$$\Pr\left[\{k\} = \text{Cons. Set}\right] = \Pr\left[\{k\} = \underset{\{j\}}{\arg\max} \frac{\sqrt{6}\sigma_u}{\pi} \ln\left(\sum_{l\in\{j\}} \exp\left(\frac{\pi}{\sqrt{6}\sigma_u}\left(\theta q_l - \overline{p}_l + \varepsilon_{il,t}\right)\right)\right) - \sum_{l\in\{k\}} C_{il,t}\right].$$

This can be rewritten as

$$\Pr\left[\{k\} = \text{Cons. Set}\right] = \Pr\left[\{k\} = \underset{\{j\}}{\arg\max} \sum_{l\in\{j\}} \exp\left(\frac{\pi}{\sqrt{6}\sigma_u}\left(\theta q_l - \overline{p}_l - \sum_{l\in\{j\}} C_{il,t} + \varepsilon_{il,t}\right)\right)\right]. \quad (12)$$

We can see that there are two intercept parameters in the probability formulation in equation (12): the true quality of the brand, q_l, and the baseline search cost, C_0. We cannot separately identify these two parameters. Lattin and Roberts (1991) encounter similar identification problem and use self-reported consumer values to get the estimates for the search cost C_0. Andrews and Srinivasan (1995) and Bronnenberg and Vanhonacker (1996) circumvent this identification problem by eliminating the intercept term in the specification of consideration utility.

However, if we allow for quality learning (as in the specification developed in section 2.2) instead of adding the statistical error term, the probability of consideration in this case would be:

$$\Pr\left[\{k\} = \text{Cons. Set}\right] = \Pr\left[\{k\} = \underset{\{j\}}{\arg\max} \sum_{l\in\{j\}} \exp\left(\frac{\pi}{\sqrt{6}\sigma_u}\left(\theta \omega_{il,t} - \overline{p}_l - \sum_{l\in\{j\}} C_{il,t}\right)\right)\right] \quad (13)$$

---

16. This is equivalent to allowing consumers to differ in their price sensitivity. Further, letting θ to be gamma distributed ensures that negative price sensitivity for all consumers.

Recall from equation (5), the mean perceived quality of brand *l* for consumer *i* on purchase occasion *t*, $\omega_{il,t}$, in equation (13) is the sum of *t* normal random variables: $\omega_{ij,t} = \sum_{l=0}^{t-1} N_{ij,l}$. Since the variance of the distribution of $\omega_{il,t}$ evolves over the consumption experience *t*, it allows us to separately identify the parameters α_0 in the variance of $\omega_{il,t}$ and the search cost $C_{il,t}$.

2.5 Preference-based versus Inertial Brand Loyalty

State dependence refers to the observed pattern in many frequently purchased product (FPP) categories wherein a consumer repeatedly buys the same brand over successive purchase occasions even in the presence of competitive price promotions. The fact that a consumer repeatedly buys the same brand could be attributed to two factors:

a) She "considers" all the brands on each purchase occasion, but continues to buy the same brand due to her high intrinsic preference for the brand. We refer to this type of repeat purchase behavior as "active loyalty"; and,

b) Due to her high search costs, she has a low propensity to search and thus only considers the brand that she repeatedly purchases. We refer to this type of repeat purchase behavior as "passive loyalty" or "inertia". Note that unlike active loyalty that is due to high intrinsic preference for a brand, in the case of passive loyalty, a consumer may repeatedly buy a brand even if she has similar preference for the competing alternatives due to "lock-in". This is because, due to her low search propensity, she fails to notice price promotions offered by other brands.

Viewed in this sense, loyalty is defined on a bounded continuum between "passive loyalty" on one end (consideration set containing one brand) and "active loyalty" on the other end (consideration set containing all the brands). Such a conceptualization is also consistent with the literature that has identified preference-based and switching cost-based rationale for repeat purchase behavior (Jeuland, 1979; Wernerfelt, 1991).

Our formulation allows us to identify the salience of these two sources as the driver of repeat purchase behavior in a product category. This distinction relies on the following result:[17]

> **PROPOSITION 1:** *Greater the time a consumer consecutively stays in the state of inertia, greater is the probability that she would add more brands into her consideration set on the next purchase occasion. In other words, a consumer does not stay in the state of inertia for long.*

17. Proofs of all the propositions are given in the Appendix.

Therefore, as the consumer repeatedly buys a brand due to inertia, the probability that she will continue to buy the same brand on the next purchase occasion successively decreases. The intuition here is that greater the time the consumer stays in a state of inertia, the probability of getting a large enough negative draw for that brand increases. If the draw becomes sufficiently negative, it will no longer be optimal for the consumer to sample that brand alone. This property sets a clear distinction between active and passive loyal consumers. A passively loyal consumer might buy the same brand for 5 or 6 time periods out of inertia ("lock-in"), but after that she is likely to include other brands in her consideration set and this might effect a brand switching.[18] Thus, in the case of a passively loyal consumer, we will observe phases of inertia followed by a brand switching that in turn is followed by another spell of inertia. On the other hand, an actively loyal consumer buys the same brand regardless of whether other brands enter their consideration set or not. Thus, she will tend to buy the preferred brand for a longer period of time.

From a managerial perspective, it is important to identify the source of state dependence. If inertia is the prime driver of repeat behavior, to induce brand switching sales promotions must be heavily supported with promotional advertising and in-store activities such as feature advertisement and displays. In this case, a shallower but well supported promotion would be more effective than a deeply discount one that is not advertised.[19]

2.6 Determinants of the Size of the Optimal Consideration Set

An insight that we obtain from the model regarding the size of the optimal consideration set is summarized in the following proposition:

> **PROPOSITION 2:** *The higher the variance of (indirect) utility, larger is the size of the optimal consideration set.*

An immediate implication of Proposition 2 is that if consumers in a product market exhibit higher price sensitivity, they will search the prevailing prices of more brands (i.e., will have larger consideration sets) before making a brand choice. This makes intuitive sense because higher price sensitivity implies that consumers will attach greater importance to discovering lower prices and hence exhibit higher search propensity. As such, the intensity of competition in these markets will be higher. We summarize this discussion in the following corollary.

---

18. Note that the traditional discrete choice model will ascribe such a switching behavior of a consumer even in the absence of any change in marketing mix variables after having bought a particular brand for a long time to the idiosyncrasies of the consumer or to unknown external market shocks.

19. In a somewhat different context, this might explain why stores such as high-end department stores such as Marshall Field's and J.C. Penny rely mostly on advertised sales while low-end stores such as T.J. Maxx and Filene's Basement have mostly unadvertised sale. The argument here is that the high-end stores cater mostly to high quality

> **COROLLARY 1**: *In product categories where consumers have lower intensity of preference for quality (equivalently, higher price sensitivity), the size of the optimal consideration set would be larger compared to product categories with lower consumer price sensitivity. Further, a consumer with higher price sensitivity would tend to search the posted prices of a larger number of brands before making a brand choice decision.*

Another implication of Proposition 2 is that if price variability is high, either because of frequent promotions or deep discounts offered on promotions, the incentive to engage in price search will be high. As such, the average size of the consideration set will be large. This has important managerial implications concerning the long-term impact of price promotions. This result demonstrates that more reliance on price promotion induces consumers to engage in more active price search. This in turn would mean more intense price competition in the future. We summarize this intuition in the following corollary.

> **COROLLARY 2**: *Greater the price variability in a product category, greater will be the size of the optimal consideration set. Thus, in product categories where promotions are frequent and are of sizable depth, consumers will actively search across brands for low prices. In contrast, in product categories where promotions are infrequent or with shallow discounts, consumers tend to repeat buy a brand out of inertia or consumer "lock-ins".*

3. DATA, ANALYSIS AND DISCUSSION

3.1 Data and Variable Definition

For model calibration and detailed analysis, we use the ERIM data set for liquid detergents. We have taken 4 national brands for the analysis: Wisk, Tide, Era and Surf. These four brands account for a total of 81% of the market share in this product category. The data set comprises a random sample of 400 households with purchase observations extending from the 25th week of 1986 to the 34th week of 1988. The minimum number of purchase observation for a household was 12; the maximum was 81 with the mean being 21 purchases. We randomly picked 200 households for the purpose of parameter estimation; the rest 200 households form the holdout sample to test the predictive validity of the proposed and competing specifications. The estimation sample had a total of 3592 purchase observations and the hold out sample had a total of 3019 purchase observations. In both the estimation and the holdout samples, we take the first five observations for each household to initialize consumers' prior beliefs about the qualities of these brands and the store familiarity variable for the consumers. The summary statistics for the entire sample of 400 households are given in Table 1A.

---

sensitive consumers with high search (shopping) costs while low-end stores primarily serve low quality sensitive but low search cost consumers.

In addition, we performed a limited cross-category analysis using ketchup data set.[20] For our analysis, we have taken 4 brands: Heinz, Hunt's, Del Monte and the Generic brand. Since we do not compare the proposed model with the competing models for ketchup data set, we do not use any hold out sample. The data set comprises of 150 randomly selected households with a total of 2332 observations. The minimum number of purchase observation for a household was 7; the maximum was 62 with the mean being 16 purchases. We use the first five observations to initialize the consumers' prior quality beliefs and the store familiarity variable. The summary statistics for the data set are given in Table 1B.

In the pre-estimation sample, we assume that consumers have the same quality beliefs for all the brands at the beginning of their purchase histories. In other words, we assume the initial values of the expected quality beliefs to be the same across all brands and across all consumers and without loss of generality set it to zero, i.e., $\omega_{ij,0} = 0 \quad \forall i, \forall j$. We then calibrate the model for the pre-estimation sample and calculate the posterior quality beliefs for all the brands and for every consumer at the end of their fifth purchase occasion. We then take these posterior quality beliefs as the initial prior beliefs for the estimation sample. The posterior quality beliefs for liquid detergent data set are given in Table 1C and for ketchup data set in Table 1D.

Factors affecting Cost of Price Search:[21] The cost of searching the posted price of a brand ($C_{ij,t}$) is the cost incurred by a consumer to ascertain the actual price of a brand on a given purchase occasion. The higher the time spent to do so, higher will be the consumer's search cost. Similarly, higher the consumer's opportunity cost of time, higher will be her search cost.

We posit that consumer *i*'s costs associated with searching the posted price of brand *j* on purchase occasion *t* depends on:

(i) Whether the brand is on display on that purchase occasion;

(ii) Whether the brand is feature advertised on that purchase occasion;

(iii) Whether the consumer is familiar with the store environment;

(iv) Whether the purchase occurs on a weekday or a weekend;

(v) Whether there is a full time homemaker in the household;

(vi) Whether the per capita household income is high or low; and,

(vii) Whether the brand was on display on the previous purchase occasion.

---

20. Liquid detergent data set has been previously used by Siddarth, Bucklin and Morrison (1995). Bronnenberg and Vanhonacker (1996) use a European data set for dry detergents. Ketchup data set has been used earlier by Andrews & Srinivasan (1995) and Chiang, Chib and Narasimhan (1999).

21. We thank the Area Editor and two anonymous reviewers for several excellent suggestions on modeling consumer search costs.

Thus, we assume that the search cost incurred by consumer i in including brand j in her consideration set on purchase occasion t is given by

$$\begin{aligned} C_{ij,t} \quad = \quad & C_0 + C_1 \times DISPLAY_{ijt} + C_2 \times FEATURE_{ijt} + C_3 \times STLOY_i \\ & + C_4 \times DOW_{it} + C_5 \times FTHM_i + C_6 \times INCOME_i + C_7 \times DISPLAY_{ij,t-1} \end{aligned} \quad (15)$$

In equation (15), $DISPLAY_{ijt}$ = 1 if brand j is displayed on purchase occasion t and = 0 otherwise. The indicator variable $FEATURE_{ijt}$ is similarly defined. Since displays and feature ads give price information (Blattberg and Neslin, 1990), they automatically reduce price search costs to zero.[22] However, we would not expect every consumer to observe displays and feature ads; thus at the aggregate level, we would only expect the search costs for ascertaining a brand's posted price to reduce by a certain fraction (and not to zero).

The time spent by a consumer to ascertain the posted price of a brand would also depend on the consumer's familiarity of its shelf location. In equation (15), we capture this through two variables viz., $STLOY_i$ and $DISPLAY_{ij,t-1}$. The variable $STLOY_i$ captures the store and category familiarity effects: the more familiar the consumer is with the store layout and the category shelf arrangement, less time consuming is the price search process (Desai and Hoyer, 2000). Following Murthi and Srinivasan (1999), it is defined as the fraction of dollars spent in the product category in the store in the initialization period. We believe that this is a much better way of capturing store and category familiarity than defining two separate variables for store and category familiarity each. $DISPLAY_{ij,t-1}$ is a dummy variable that indicates whether brand j was on display on previous purchase occasion t-1. The argument here is that that since displays are usually placed on special locations, if on the previous purchase occasion the brand was on display, it would make it more difficult for the consumer to search for the brand if it is not displayed in the current period.

We capture the impact of opportunity cost of time on search costs through three variables viz. DOW_{it}, $FTHM_i$, and $INCOME_i$. DOW_{it} is an indicator variable such that DOW_{it} = 1 if consumer i undertakes the t^{th} purchase during a weekend and = 0 if the purchase was made during a weekday. It captures the shopping time available to the consumer. On weekends, since working households have more time available for shopping, we would expect them to engage in more extensive price search due to a lower opportunity cost of time. $FTHM_i$ is an indicator variable such that $FTHM_i$ = 1 if household i has a full-time homemaker and = 0 otherwise. We would expect the presence of a full time homemaker to reduce the household's opportunity cost of time and hence to reduce price search costs. $INCOME_i$ is

22. In the liquid detergent category, brands are often feature advertised and displayed. The frequency of feature advertisements varies from 9% to 14% across the four selected brands. Similarly, the frequency of displays varies from 11% to 22% across the brands.

defined as the per capita household income and is measured in dollars. The argument here is that a high-income consumer would have a higher opportunity cost of time and hence a higher search cost for comparing prices across brands in a product category.

Based on the above discussion, we would expect the signs of the parameters to be: $C_0 > 0$; $C_1 < 0$; $C_2 < 0$; $C_3 < 0$; $C_4 < 0$; $C_5 < 0$; $C_6 > 0$; and, $C_7 > 0$.[23]

3.2 Parameter Estimates and Comparison with Competing Models

We used the Method of Simulated Moments (MSM), proposed by McFadden (1989) and Pakes and Pollard (1989), to estimate the parameters of the proposed model and three other competing models. We coded the program in MATLAB. We estimated the following parameters for the proposed model in this analysis:

(i) Mean quality sensitivity parameter, $\bar{\theta}$;

(ii) Variance of quality sensitivity across the population, σ_θ^2. We assume quality sensitivity parameter θ to be gamma distributed with mean $\bar{\theta}$ and variance σ_θ^2;

(iii) Ratio of the noise in consumption signal to the information that can be gained at the beginning of the observation period, α_0;

(iv) The search cost parameters viz.,

(a) C_0 – Baseline search cost per brand for a household with per capita income of $10,000 in the absence of feature advertisement and displays; the household has no full time homemaker; the store-category familiarity for the household is zero and the purchase is made during the weekdays;

(b) C_1 – Effect of display on search costs;

(c) C_2 – Effect of feature advertisement on search cost;

(d) C_3 – Effect of store-category familiarity on search costs;

(e). C_4 – Effect of day of the week (when the brand was purchased) on search costs;

(f) C_5 – Effect of presence of a full time homemaker on search costs;

(g) C_6 – Effect of increase in per capita household income by $1,000 on search costs;

(h) C_7 – Effect of displays on the previous purchase occasion on current search costs.

---

23. These represent our hypotheses about the impact of various factors on the intensity of price search behavior. The sign of the parameter is ultimately determined by the data.

The parameter estimates for the liquid detergent data set are given in Table 2A. The parameters are by and large statistically significant and in the anticipated direction.

For this data set, we compared the predictive power and goodness-of-fit of the proposed specification (MODEL I) to four other competing specifications:

(i) A nested model (MODEL II) obtained by ignoring the consideration stage. The implied parametric restrictions are $C_0 = C_1 = C_2 = C_3 = C_4 = C_5 = C_6 = C_7 = 0$. In this specification, it is assumed that consumers consider (i.e. are aware of the posted prices of) all the four brands on each purchase occasion. The model, however, allows for quality learning. To ensure that Models I and II have the same set of variables, we assume that displays and feature advertisements affect preferences or quality beliefs of the competing brands.[24] The parameters that we estimate in the nested model are

a) Mean quality sensitivity parameter, $\bar{\theta}$;

b) Variance of quality sensitivity across the population, σ_θ^2

c) Ratio of the noise in consumption signal to the information that can be gained at the beginning of the observation period in the estimation sample, α_0;

d) Ratio of the noise in consumption signal to the noise in feature advertisement, κ_f;

e) Ratio of the noise in consumption signal to the noise in displays, κ_d;

The parameter estimates for Model II are given in Table 2A.

(ii) The model proposed by Andrews and Srinivasan (1995) – MODEL III – that models consideration and choice stages in a reduced-form framework. The choice and consideration utilities are assumed to be different; a consumer considers only those brands whose consideration utility exceeds a certain threshold. The consideration and choice utilities of a brand are functions of price, display, feature ads and brand loyalty. The parameter estimates for Model III are summarized in Table 2B.

(iii) The model proposed by Siddarth, Bucklin and Morrison (1995) – MODEL IV – that again models consideration and choice stages in a reduced-form framework. In this specification, a consumer considers only those brands whose fraction of the past purchases exceeded a certain threshold. The choice utility of a brand is a function of price, display, feature ads and brand loyalty. The parameter estimates for the model are summarized in Table 2C.

(iv) The model proposed by Bronnenberg and Vanhonacker (1996) – MODEL V – that models consideration and choice stages in a reduced-form framework. At the consideration stage, only the

24. Technical details about quality evolution in the presence of quality signals from consumption experience, display and feature ads are given in the Technical Appendix.

salience features of a brand – viz., recency of purchase, display and feature ads, and price range membership – play a role. At the choice stage, only intrinsic preference and posted price play a role. Similar to Andrews and Srinivasan (1995), a consumer considers only those brands whose consideration utility exceeds a certain threshold. Consumer heterogeneity is captured through finite discrete segments (Kamakura and Russell, 1989). The parameter estimates for the model are summarized in Table 2D.

We do not compare the proposed specification with the framework developed in Lattin and Roberts (1991) since one needs self-reported consumer data to estimate their model. Further, we do not compare the proposed model with Chiang, Chib and Narasimhan (1999) since it requires hierarchical Bayesian estimation methods while we employ the MSM estimation approach.

Table 3B summarizes the results from comparison of the proposed model (MODEL I) against the four competing models (MODEL II-V) using the statistical test proposed by Singleton (1985) that is appropriate for comparing non-nested specifications in the Generalized Method of Moments (GMM) framework.[25] We report the test results for both the estimation and the holdout samples. We find that the test rejects MODEL II in favor of MODEL I for both the estimation and the holdout samples (p-value< 0.01). This supports the inference that while making a brand choice decision, consumers "consider" only a subset of brands on any purchase occasion. Similarly, the test rejects Models III and IV for both the estimation and the holdout samples (p-value<0.1). This allows us to infer that relative to the competing Models III and IV, the proposed Model I is a superior representation of consumer brand choice behavior for liquid detergents. While the test fails to reject Model V for the estimation sample (p-value>0.10), it rejects it in favor of Model I for the holdout sample. Thus, based on the Singleton test (1985), we conclude that Model I performs better the models of consideration set proposed by Andrews and Srinivasan (1995), Siddarth, Bucklin and Morrison (1995) and Bronnenberg and Vanhonacker (1996).

Table 3A reports two conventional measures of goodness-of-fit for the holdout sample: Hit Rate and Log Likelihood value. On both these criteria, we find that the proposed model outperforms the competing specifications. In fact, we find that for the holdout sample, Models II, III and V predict brand choice quite poorly; while Model IV does a decent job, it does not predict choices as well as Model I.

---

25. In the Singleton test (1985), the null hypothesis is that the competing model is the true model of consumer behavior. The test thus entails whether the proposed model (i.e., Model I) significantly outperforms the competing model. The test statistic is distributed χ^2 with 1 d.f.

3.3 Results for Liquid Detergents and Discussion of Managerial Insights

We summarize our discussion under five broad substantive issues:

Perceived Quality Positioning of the Competing Brands: From Table 2A, we find the estimated mean perceived quality of Wisk, Tide, Era and Surf to be 0.0097, 0.1729, 0.1135 and 0.0589, respectively. This is somewhat surprising since Wisk happens to be the market share leader (31% share). As expected, the brand with the lowest market share viz., Surf (20% share) has the lowest perceived quality. It thus appears that several consumers are purchasing Wisk simply because they do not consider any other brand. This is also consistent with the fact that Wisk is the most feature-advertised (frequency of feature ads = 0.14) and displayed (frequency of display = 0.21) brand.

Consumer Learning and the Effectiveness of Quality Cues: The estimate of α_0 – the initial value of precision of quality beliefs – is 2.6758 (assumed to be the same across all consumers and across all brands). Recall that a high value of α_0 suggests a low variance in subjective quality beliefs held by the consumers at the beginning of their consumption histories. Thus a value of α_0 implies that the extent of quality learning that can occur through consumption experience is low. Our estimate of α_0 suggests a moderate extent of quality learning happening in the product category. Figure 1 shows the evolution of quality beliefs for the four brands during the first 15 purchase occasions. We observe that even after 15 consumption experiences with the brand, the consumer is still updating her quality beliefs. For the nested model (Model II), the estimate of α_0 is 1.4008. This suggests that if we ignore consideration effects, we may overestimate the extent of quality learning happening in a product category. This makes sense because if we ignore heterogeneity in consideration set sizes (across consumers and over purchase occasions), all the variance is attributed to heterogeneity in subjective quality beliefs thereby magnifying consumer-learning effects. As noted in section 2.3, heterogeneity in subjective quality beliefs in the proposed specification is akin to intercept heterogeneity in the reduced-form specifications such as the Logit model (e.g., Chintagunta, Jain and Vilcassim, 1991). Thus, we conclude that ignoring consideration effects may lead to overestimation of consumer heterogeneity, thereby confirming the conclusions in Chiang, Chib and Narasimhan (1999).

Recall that to ensure comparability of Models I and II, in Model II we allow displays and feature ads to affect intrinsic brand preferences (as in Erdem and Keane, 1996). This is captured by κ_f (the ratio of the noise in consumption signal to the noise in feature ad) and by κ_d (the ratio of the noise in consumption signal to the noise in display), respectively. High values of κ_f and κ_d imply that feature ads and displays have a strong impact on intrinsic preference. The parameter estimates of κ_d = 0.0041 and κ_f = 0.0062 suggest almost no impact of feature ads and displays on quality beliefs. This is expected since feature ads and displays primarily give price information (Blattberg and Neslin, 1990).

<u>Quality-Price Trade-off and Consumer Price Sensitivity</u>: The estimates of the mean quality sensitivity parameter, $\bar{\theta}$, for Model I and II are 2.4168 and 4.7522, respectively. This suggests that ignoring consideration effects might lead to underestimation of price sensitivity (high quality sensitivity or a high marginal willingness-to-pay for quality is synonymous to low price sensitivity). This makes sense because in Model II, we assume that a consumer observes the prices of all the brands and hence whenever any brand is on sale, the consumer notices such sales promotion and responds to it. In contrast, Model I explicitly recognizes that consumers may not be aware of the posted prices of all the brands and hence a non-response to ongoing price promotion need not necessarily imply low price sensitivity.

These insights are further reinforced by comparing the price elasticities implied by Models I and II (Tables 6A-6C). As expected, we find that the conditional price elasticities (conditional on consideration) are greater for Model I than price elasticities for Model II. However, we cannot make a general statement about the unconditional price elasticites for Model I. In particular, we note that the unconditional price elasticities from Model I are higher than the price elasticities from Model II only for the high market share brands (viz., Wisk and Tide). In contrast, the unconditional price elasticities for Model 1 are lower than the price elasticities for Model II for the low market share brands (viz., Era and Surf).[26]

For Model III, the estimated price sensitivity in the consideration stage is statistically insignificant. This suggests that posted prices may not have any appreciable impact on consideration probability. This finding supports the contention of Bronnenberg and Vanhonacker (1996) that it is the price range membership (high- vs. low-price tier) rather than the actual posted prices that determine consideration set composition. The estimates of the variance of quality sensitivity, σ_θ^2, from Model I and II are 0.4251 and 1.219, respectively. This again highlights the fact that ignoring heterogeneity in consideration sets may lead to overestimation of consumer heterogeneity as noted by Chiang, Chib and Narasimhan (1999).

<u>Determinants of the Costs of Price Search</u>: The estimate for the baseline search cost C_0 is 0.0472. Thus, the average size of the consideration set in the absence of any marketing activity for a consumer with no store familiarity and who comes from a household with per capita income of $10,000 and with no full time homemaker is 2.31. Such a consumer considers only one brand during 23.4% of the purchase occasions, two brands during 32% of the purchase occasions, three brands during 35.3% of the purchase occasions and four brands for the rest 9.1% of the occasions (Table 4). This implies that a consumer is

26. We thank the Area Editor and an anonymous reviewer for alerting us to this distinction.

likely to repeatedly buy a brand out of inertia, thereby suggesting significant degree of “passive loyalty” in this product category.

Table 4 also summarizes the marginal impact of feature ads and displays on the size of the consideration set. The reduction in search cost due to display, C_1, is - 0.0091. Thus, if all the brands were on display, the average size of consideration set increases from 2.31 to 2.50. The reduction in search cost due to feature advertisements, C_2, is - 0.0197. If all the brands were to be feature advertised, then the average size of consideration set increases from the base line case of 2.31 to 2.75. This suggests that displays are not as effective as feature ads in influencing search behavior. This is an interesting finding: in this product category, except for Era, the other three brands viz., Wisk, Tide and Surf seem to rely more heavily on in-store display than feature ads. This suggests that these brands need to rethink their promotional-mix (unless otherwise justified by cost considerations).

The reduction in search cost due to store familiarity, C_3, is - 0.002 and is not significant even at α=20%. The reduction in search cost during shopping trips on weekend as compared to those during weekdays, C_4, is - 0.0032 and is significant at α=10%. Thus, the average size of consideration set during weekends is 2.48 while during the weekdays is 2.43 thus suggesting a very modest impact of DOW on search behavior. This is contrary to our expectation that consumers engage in more extensive price search during weekends because they have more time available for shopping (low opportunity cost of time). We conjecture that this could be due to the fact that stores are more crowded during weekends that might dissuade consumers from searching extensively.

Income has a significant impact on search costs. The increase in search cost for every $1,000 increase in per capita household income, C_6, is 0.0018. In the data set, per capita household income varies from $1,000 to $33,000. Thus, the average consideration set size for a household with per capita income of $10,000 is 2.31. The average consideration set size corresponding to per capita income of $20,000 and $30,000 are 1.99 and 1.83, respectively. This shows that the intensity of price search decreases with income.

The increase in search cost due to previous display, C_7, is 0.0009 and is not significant even at α=20%. Thus, previous displays do not significantly impact the costs of search during the current purchase occasion.

Consumer Search Behavior: Table 5A reports the frequency of the consideration set sizes for the proposed model when each of the four brands viz., Wisk, Tide, Era and Surf is chosen. We observe that the average size of the consideration set when Wisk was bought is 2.21; for Tide the average size is 2.42; for Era the average size is 2.27; and, for Surf the average size is 2.35. Thus, it appears that in spite of its lower perceived quality, Wisk enjoys market share leadership primarily due to the fact that several consumers do not consider other brands.

To further investigate the extent of active versus passive loyalty being enjoyed by the four brands, we compared the average search costs for the households who are loyal towards these brands. We divided the consumer population into five segments: Wisk-loyal, Tide-loyal, Era-loyal, Surf-loyal, and switchers. We classified a consumer as loyal to a brand if she bought the brand on more than 50% of the purchase occasions. We find that the average per capita income for Wisk-loyal consumers is \$11,000, for Tide-loyal consumers is \$9,700 for Era-loyal consumers is \$10,000 and for Surf-loyal consumers is \$5,600. Since search costs increase with income, this shows that Wisk enjoys the highest degree of passive loyalty while Surf has the highest degree of active loyalty.

Overall, we find that in the liquid detergents category, consumers consider one brand 26.2% of the times, two brands 28.7% of the times, three brands around 33.1% of the times and four brands the rest 12% of the times. This corresponds to an average consideration set size of 2.31. This indicates moderate intensity of price search in the product category. Table 5B compares the set size probabilities predicted by the proposed model and the competing models III to V. The average consideration set size for Models III, IV and V are 2.494, 2.564 and 2.585, respectively. We observe that the proposed model predicts a smaller consideration set. The reason we get lower consideration set size is as follows. The models proposed in Andrews and Srinivasan (1995), Siddarth, Bucklin and Morrison (1995) and Bronnenberg and Vanhonacker (1996) ignore the brand intercept term in the consideration stage. This is because in their reduced-form formulations, brand intercept is not separately identifiable. If we include intrinsic brand preference (through intercept) at the consideration stage, then some brands have a higher probability of entering the consideration stage. This would lead to smaller consideration set sizes comprising brands having large brand equities.[27] Note that if we ignore intrinsic brand preference, we would get biased estimates of the distribution of consideration set sizes.

3.4 Cross-Category Analysis – Price Search Behavior in Ketchup Category

The parameter estimates for ketchup data set are given in Table 7. The parameters are by and large statistically significant and in the anticipated direction.

The estimate of mean quality sensitivity, $\overline{\theta}$, is 3.2140. This suggests that consumers are more prices sensitive while purchasing liquid detergents than while purchasing ketchup. This is to be expected

---

27. For the proposed model, we recomputed the distribution of consideration set sizes setting intercept term, $C_0 = 0$. We find that in this case consumers consider one brand 19.7% of the times, two brands 27.9% of the times, three brands around 30.1% of the times and four brands the rest 22.3% of the times. Thus, the predicted average size of consideration set is 2.55 which is comparable to those from Models III-V.

since detergents are a higher expense item than ketchup. This may also suggest that through advertising, the leading brands in this category have been able to increase the intensity of preference for the differentiating attribute thereby reducing price sensitivity. We find the variance of quality sensitivity, σ_θ^2 = 0.4739 which is comparable to that for liquid detergents.

The value of the intrinsic search cost C_0 is estimated at 0.0284 and is lower than that for liquid detergents. This suggests that price search is less costly for ketchup. We conjecture that this could be due to the fact that shelf space assigned to ketchup category is much smaller than that assigned to liquid detergents, thereby making price comparisons easier for ketchup. The reduction in search costs due to feature ads, C_2, is - 0.0142. Thus, if all the brands are feature advertised, the average consideration set size increases from 2.65 to 2.75. Further, a lower value of C_2 for ketchup relative to that for liquid detergents suggests that consumers are more likely to see feature ads for liquid detergents than for ketchup. The reduction in search costs due to displays, C_1, is - 0.0210 and is significant. In fact, we find that for ketchup displays are more effective than feature ads. This is in contrast to liquid detergents where feature ads are more effective than displays in reducing search cost. This again is an interesting finding: all the brands in this category seem to rely more heavily on feature ads than in-store display! This suggests that these brands need to rethink their promotional-mix (unless otherwise justified by cost considerations).

Similar to liquid detergents, in ketchup category we find no significant impact of store familiarity and timing of purchase (weekend/weekday). Additionally, unlike liquid detergents, we find no significant impact of income on search cost.

Table 8 reports the frequency of consideration set sizes for the proposed model when each of the four brands viz. Heinz, Hunt’s, Del Monte and Generic is chosen. We observe that the average size of the consideration set in when Heinz was bought is 2.125; for Hunt’s the average size is 2.536; for Del Monte the average size is 2.541; and, for the Generic brand the average size is 2.819. As expected, consumers who buy the Generic brand -- being more price sensitive – undertake the most extensive search. As in the liquid detergent category, we find that the market share leader, Heinz, gains from passive loyalty.

Overall, we find that in the ketchup category, consumers consider one brand 28.7% of the times, two brands 27.6% of the times, three brands around 30.9% of the times and four brands the rest 12.9% of the times. This corresponds to an average consideration set size of 2.47. This indicates that consideration sets are larger for ketchup than for liquid detergents.

4. CONCLUSIONS

In this paper, we advance a structural model of price search wherein a consumer engages in an optimal trade-off between incurring costly search against the additional potential benefits arising from price search while deciding on the number of brands to search. This is motivated by the fact that due to frequent price promotions of varying depths of discount, there exists considerable price uncertainty; while consumers are aware of the distribution of prices, they do not know the actual posted price of a brand on a given purchase occasion unless they search. We conceptualize this price search as the basis for consideration set effects. Specifically, in our conceptualization, the set of brands that the consumer selects to sample posted prices of on any particular purchase occasion is termed as the consumer's optimal consideration set. After making the optimal consideration set choice, the consumer selects the brand (from among those in her consideration set) that yields the highest expected surplus. Our structural analysis of consideration set effects reiterates several important behavioral aspects of brand choice.

The proposed specification allows us to distinguish between two rationale for brand loyalty: (i) High search cost underlying inertia or "passive loyalty"; and, (ii) High intrinsic preference underlying "active loyalty". Passive loyalty refers to the case when a consumer repeatedly selects the same brand because her high search costs does not outweigh the potential benefits from including more brands in her consideration set. In contrast, an actively loyal consumer searches more than one brand (since her search costs are low); however, she still repeat-buys the same brand since her intrinsic preference for the brand is high. Our empirical findings suggest that marketing activities, primarily feature advertisement, play a critical role in increasing consumers' search propensity. We find that while feature advertising impacts price search behavior more than displays for liquid detergents, the opposite holds true for ketchup. Interesting, in both these categories we find that brands are using inappropriate promotional mix (unless otherwise justified by cost considerations).

We consider this paper as an important first step to study consideration effects in the context of FPPs from a structural modeling perspective using scanner panel data. Having said that, we realize the limitations of the proposed econometric specification. For instance, we assume that the consumer adopts the fixed sample search strategy (Stigler, 1961) in discovering the posted prices of the brands in her consideration set. It may be interesting to investigate the insights about consideration effects using a sequential search (Roberts and Stahl, 1993) framework. Further, while we recognize consumer uncertainty about brand qualities and model evolution of consumer quality perceptions, since our model is based on a static utility-maximization paradigm, it fails to capture the richness of variety-seeking behavior that can be obtained in a dynamic utility maximization framework (the so-called multi-armed bandit problem). We hope to address these issues in future research effort.

TABLE 1A: Descriptive Statistics for Liquid Detergent Data Set

Brand	Market Share	Frequency of Feature Advertisement	Frequency of Display	Mean Price (Std. Dev.)
Wisk	0.31	0.14	0.22	5.06 (0.58)
Tide	0.28	0.13	0.21	5.97 (0.61)
Era	0.21	0.07	0.06	5.92 (0.57)
Surf	0.20	0.09	0.11	5.46 (0.61)

NOTE: Price is in cents per ounce.

TABLE 1B: Descriptive Statistics for Ketchup Data Set

Brand	Market Share	Frequency of Feature Advertisement	Frequency of Display	Mean Price (Std. Dev.)
Heinz	0.68	0.37	0.07	4.25 (0.80)
Hunts'	0.16	0.10	0.02	3.48 (0.42)
Del Monte	0.08	0.09	0.02	3.81 (0.56)
Generic Brand	0.08	0.04	0.01	3.34 (0.31)

NOTE: Price is in cents per ounce.

TABLE 1C: Posterior Quality Beliefs of Brands from the Pre-Estimation Sample[28] for Liquid Detergent Data Set

Mean of the Posterior Density	MODEL I	MODEL II
ω_{Wisk} (Mean of the Posterior Quality Belief of Wisk)	- 0.0047	0.0226
ω_{Tide} (Mean of the Posterior Quality Belief of Tide)	0.0956	0.0835
ω_{Era} (Mean of the Posterior Quality Belief of Era)	0.0564	0.0430
ω_{Surf} (Mean of the Posterior Quality belief of Surf)	0.0434	0.0261

28. These quality beliefs obtained from the initialization sample are used as prior beliefs for the first purchase observation for the household in the estimation sample.

TABLE 1D: Posterior Quality Beliefs of Brands from the Pre-Estimation Sample[29] for Ketchup Data Set

Mean of the Posterior Density	MODEL I
ω_{Heinz} (Mean of the Posterior Quality Belief of Heinz)	0.1010
$\omega_{Hunts'}$ (Mean of the Posterior Quality Belief of Hunts')	0.0144
$\omega_{Del\ Monte}$ (Mean of the Posterior Quality Belief of Del Monte)	0.0149
$\omega_{Generic}$ (Mean of the Posterior Quality Belief of the Generic Brand)	0.0028

TABLE 2A: Parameter Estimates for Models I & II for Liquid Detergent Data Set

Parameter	Explanation	MODEL I (Std. Deviation)	MODEL II (Std. Deviation)
$\bar{\theta}$	Mean value of the intensity of preference (willingness-to-pay) for quality across the consumers	2.4168 (0.130)	4.7522 (0.141)
σ_θ^2	Variance of the intensity of preference (willingness-to-pay) for quality across the consumers	0.4251 (0.271)	1.2199 (0.254)
κ_f	Ratio of the informativeness (about the true quality) of Feature Advertisement Signal to that of Consumption Signal	-	0.062 (0.151)
κ_d	Ratio of the informativeness (about the true quality) of Displays to that of Consumption Signal	-	0.041 (0.010)
α_0	Inverse of the uncertainty in quality of at the beginning of consumption history; assumed same $\forall i, \forall j$	2.6758 (0.710)	1.4008 (0.671)
C_0	Baseline Search Costs for discovering the posted price of a brand	0.0472 (0.010)	-
C_1	Effect of Display on Search Costs	- 0.0091 (0.002)	-
C_2	Effect of Feature Ad on Search Costs	- 0.0197 (0.006)	-
C_3	Effect of Store-Category Familiarity on Search Costs	- 0.0020 (0.003)	-

29. These quality beliefs obtained from the initialization sample are used as prior beliefs for the first purchase observation for the household in the estimation sample.

TABLE 2A: Parameter Estimates for Models I & II for Liquid Detergents (cont.)

Parameter	Explanation	MODEL I (Std. Deviation)	MODEL II (Std. Deviation)
C_4	Effect on Search Costs if purchase was made during weekend	- 0.0032 (0.0019)	-
C_5	Effect of presence of full time homemaker on Search Costs	- 0.0002 (0.003)	-
C_6	Effect of increase in per capita household income by $1000 on Search Costs	0.0018 (0.0007)	-
C_7	Effect of Display on previous purchase occasion on Search Costs	0.0009 (0.0012)	-
$\overline{\omega}_{Wisk}$	Mean value of perceived quality of Wisk (across all consumers and across all purchase occasions)	0.0097 (0.006)	0.0431 (0.007)
$\overline{\omega}_{Tide}$	Mean value of perceived quality of Tide (across all consumers and across all purchase occasions)	0.1729 (0.008)	0.0624 (0.008)
$\overline{\omega}_{Era}$	Mean value of perceived quality of Era (across all consumers and across all purchase occasions)	0.1135 (0.015)	0.0469 (0.014)
$\overline{\omega}_{Surf}$	Mean value of perceived quality of Surf (across all consumers and across all purchase occasions)	0.0589 (0.003)	0.0307 (0.003)

TABLE 2B: Parameter Estimates for Model III for Liquid Detergent Data Set

Parameter	Explanation	MODEL III (Std. Deviation)
U_λ	Threshold Utility for Consideration	0.6241 (0.191)
ω_{Tide}	Intercept for Tide (ω_{Wisk} = 0 for identification)	0.9918 (0.001)
ω_{Era}	Intercept for Era	0.7470 (0.002)
ω_{Surf}	Intercept for Surf	0.5899 (0.002)
P_{co}	Effect of Price on Probability of Consideration	2.0204 (1.316)
A_{co}	Effect of Feature Advertisement on Probability of Consideration	1.0589 (0.487)

TABLE 2B: Parameter Estimates for Model III for Liquid Detergents (cont.)

Parameter	Explanation	MODEL III (Std. Deviation)
D_{co}	Effect of Display on Probability of Consideration	1.4722 (0.440)
L_{co}	Effect of Brand Loyalty on Probability of Consideration	2.6065 (0.052)
P_{ch}	Effect of Price on Brand Choice Probability (conditional on consideration)	- 1.8629 (0.003)
A_{ch}	Effect of Feature Ad on Brand Choice Probability (conditional on consideration)	0.4259 (0.401)
D_{ch}	Effect of Display on Brand Choice Probability (conditional on consideration)	0.5437 (0.212)
L_{ch}	Effect of Brand Loyalty on Brand Choice Probability (conditional on consideration)	2.8252 (0.156)

TABLE 2C: Parameter Estimates for Model IV for Liquid Detergent Data Set

Parameter	Explanation	MODEL IV (Std. Deviation)
λ	Cut-off Threshold for Past Purchases	0.2331 (0.0010)
θ	Mixing probability for the model in which any brand that is on promotion enters the Consideration Set	0.4106 (0.0700)
γ	Mixing probability for the model wherein there is no consideration	0.1727 (0.0401)
ω_{Tide}	Intercept for Tide (ω_{Wisk} = 0 for identification)	0.0110 (0.0021)
ω_{Era}	Intercept for Era	- 0.0005 (0.0003)
ω_{Surf}	Intercept for Surf	0.0102 (0.0010)
P_{ch}	Effect of Price on Brand Choice Probability (conditional on consideration)	- 0.9022 (0.0322)
A_{ch}	Effect of Feature Ad on Brand Choice Probability (conditional on consideration)	0.7963 (1.3751)
D_{ch}	Effect of Display on Brand Choice Probability (conditional on consideration)	0.8136 (0.5161)
L_{ch}	Effect of Brand Loyalty on Brand Choice Probability (conditional on consideration)	0.5849 (0.0565)

TABLE 2D: Parameter Estimates for Model V for Liquid Detergent Data Set

Parameter	Explanation	MODEL IV (Std. Deviation)
ω_{Tide}	Intercept for Tide (ω_{Wisk} = 0 for identification)	- 0.0076 (0.0001)
ω_{Era}	Intercept for Era	0.0059 (0.0001)
ω_{Surf}	Intercept for Surf	- 0.0055 (0.0001)
P_{ch1}	Effect of Price on Brand Choice Probability (conditional on consideration) for Segment 1	- 4.3727 (0.03)
A_{co1}	Effect of Feature Advertisement on Probability of Consideration for Segment 1	0.7013 (0.656)
D_{co1}	Effect of Display on Probability of Consideration for Segment 1	0.4378 (0.241)
L_{co1}	Effect of Brand Loyalty on Probability of Consideration for Segment 1	7.8750 (0.11)
$U_{\lambda 1}$	Threshold Utility for Consideration for Segment 1	2.2669 (0.67)
P_{ch2}	Effect of Price on Brand Choice Probability (conditional on consideration) for Segment 2	2.5266 (0.03)
A_{co2}	Effect of Feature Advertisement on Probability of Consideration for Segment 1	0.8290 (0.601)
D_{co2}	Effect of Display on Probability of Consideration for Segment 1	1.9781 (0.201)
L_{co2}	Effect of Brand Loyalty on Probability of Consideration for Segment 2	6.9023 (0.460)
λ	Relative sizes of the 2 segments [Segment 1's relative size = $e^{\lambda}/(1+ e^{\lambda})$]	1.055 (0.410)

TABLE 3A: Hit rates and Likelihood values for Models I - V for Hold out sample for Liquid Detergents

	MODEL I	MODEL II	MODEL III	MODEL IV	MODEL V
Hit Rate	71.76%	63.33%	58.01%	67.66%	62.13%
Log Likelihood	1890	2478	2644	2093	2529

TABLE 3B: Result from comparison of Proposed Model (MODEL I) with Competing Models for Liquid Detergent (MODEL II-V)

Model Comparison using Singleton Test	Singleton Test Results	
	Estimation Sample	Hold-out Sample
MODEL I against MODEL II H_0 : MODEL II is the "true" model H_1 : MODEL I is the "true" model	J_S = 5.91(χ^2 with 1 d.f.) p-value < 0.01	J_S = 6.84 (χ^2 with 1 d.f.) p-value < 0.01
MODEL I against MODEL III H_0 : MODEL III is the "true" model H_1 : MODEL I is the "true" model	J_S = 3.67 (χ^2 with 1 d.f.) p-value <0.10	J_S = 7.93 (χ^2 with 1 d.f.) p-value < 0.01
MODEL I against MODEL IV H_0 : MODEL IV is the "true" model H_1 : MODEL I is the "true" model	J_S = 3.73 (χ^2 with 1 d.f.) p-value < 0.10	J_S = 3.94 (χ^2 with 1 d.f.) p-value <0.10
MODEL I against MODEL V H_0 : MODEL V is the "true" model H_1 : MODEL I is the "true" model	J_S = 2.53 (χ^2 with 1 d.f.) p-value > 0.1	J_S = 10.51 (χ^2 with 1 d.f.) p-value <0.01

TABLE 4: Predicted Probabilities for Consideration Set Size

With varying Search Costs

Specification for Price Search Costs	Size of the Consideration Set			
	Pr[# Brands = 1] (Std. Deviation)	Pr[# Brands = 2] (Std. Deviation)	Pr[# Brands = 3] (Std. Deviation)	Pr[# Brands = 4] (Std. Deviation)
$C_{ijt} = C_0$ [30]	0.234	0.320	0.353	0.091
$C_{ijt} = C_0 + C_1 \times DISPLAY_{ijt}$ (Display for all brands)	0.212	0.273	0.315	0.198
$C_{ijt} = C_0 + C_2 \times FEATURE_{ijt}$ (Feature ad for all brands)	0.182	0.214	0.266	0.336
$C_{ijt} = C_0 + C_3 \times STLOY_i$ (Store Loyalty = 1)	0.229	0.310	0.356	0.103
$C_{ijt} = C_0 + C_4 \times DOW_{it}$ (Purchase is made on a weekend, DOW=1)	0.227	0.303	0.360	0.108
$C_{ijt} = C_0 + C_6 \times INCOME_i$ (Per capita household income of $20,000)	0.289	0.473	0.187	0.048
$C_{ijt} = C_0 + C_6 \times INCOME_i$ (Per capita household income of $30,000)	0.341	0.503	0.132	0.022
$C_{ijt} = 0$ (Price search is cost-less)	0 (By assumption)	0 (By assumption)	0 (By assumption)	1 (By assumption)

30. Baseline search cost when there is no feature ad and no display for a consumer with store-category familiarity = 0; purchase is made on a weekday; per capita household income = $10,000; full-time homemaker = 0.

TABLE 5A: Consideration Set Sizes when the respective Brands are chosen

(Liquid Detergent Set)

Consideration Set Probabilities	Pr[# Brands = 1]	Pr[# Brands = 2]	Pr[# Brands = 3]	Pr[# Brands = 4]
Wisk chosen	0.322	0.288	0.246	0.144
Tide chosen	0.241	0.263	0.331	0.165
Era chosen	0.252	0.278	0.421	0.049
Surf chosen	0.208	0.329	0.367	0.096

TABLE 5B: Consideration Set Sizes for Models I, III, IV and V for Liquid Detergents

Consideration Set Probabilities	Model I	Model III	Model IV	Model V
Pr[# Brands = 1]	0.262	0.209	0.192	0.197
Pr[# Brands = 2]	0.287	0.266	0.254	0.240
Pr[# Brands = 3]	0.331	0.347	0.352	0.344
Pr[# Brands = 4]	0.120	0.178	0.202	0.219

TABLE 6A: Model I – Conditional Price Elasticities[31] for Liquid Detergents

Conditional Price Elasticity	Wisk Market Share	Tide Market Share	Era Market Share	Surf Market Share
Price of Wisk	- 2.057	1.122	1.191	1.845
Price of Tide	0.536	- 2.123	0.592	0.540
Price of Era	0.539	0.397	- 2.667	0.924
Price of Surf	0.879	0.882	0.924	- 2.927

TABLE 6B: Model I – Unconditional Price Elasticities for Liquid Detergents

Unconditional Price Elasticity	Wisk Market Share	Tide Market Share	Era Market Share	Surf Market Share
Price of Wisk	- 1.360	0.155	0.052	0.264
Price of Tide	0.397	- 0.557	0.308	0.242
Price of Era	0.581	0.043	- 0.218	0.346
Price of Surf	1.588	0.425	0.072	- 1.028

31. This refers to the marginal impact of price change on brand choice probability, conditional on the fact that the brand is included in the consideration set.

TABLE 6C: Model II – Price Elasticities for Liquid Detergents

Price Elasticity	Wisk Market Share	Tide Market Share	Era Market Share	Surf Market Share
Price of Wisk	- 1.285	0.439	0.574	1.167
Price of Tide	0.272	- 1.231	0.431	0.462
Price of Era	0.267	0.020	- 1.430	0.352
Price of Surf	0.571	0.418	0.459	- 1.490

TABLE 7: Parameter Estimates for Model I for Ketchup Data Set

Parameter	Explanation	MODEL I (Std. Deviation)
$\bar{\theta}$	Mean value of quality sensitivity across the consumers	3.2140 (0.25)
σ_{θ}^{2}	Variance of quality sensitivity across the consumers	0.4739 (0.30)
α_0	Inverse of the uncertainty in quality of at the beginning of consumption history; assumed same $\forall i$, $\forall j$	3.5534 (1.10)
C_0	Baseline Search Costs for discovering the posted price of a brand	0.0284 (0.004)
C_1	Effect of Display on Search Costs	- 0.0105 (0.005)
C_2	Effect of Feature Ad on Search Costs	- 0.0142 (0.007)
C_3	Effect of Store-Category Familiarity on Search Costs	- 0.0001 (0.033)
C_4	Effect on Search Costs if the purchase was made during the weekend	- 0.0002 (0.01)
C_5	Effect of absence of full time homemaker in household on Search Costs	0.0001 (0.02)
C_6	Effect of increase in per capita household income of $1000 on Search Costs	0.0003 (0.01)
$\bar{\omega}_{Heinz}$	Mean value of the perceived quality of Heinz (across all consumers and across all purchase occasions)	0.1595 (0.006)
$\bar{\omega}_{Hunt's}$	Mean value of the perceived quality of Hunt’s	0.0199 (0.002)
$\bar{\omega}_{Del\ Monte}$	Mean value of the perceived quality of Del Monte	0.0291 (0.007)
$\bar{\omega}_{Generic}$	Mean value of the perceived quality of the Generic brand	0.0092 (0.001)

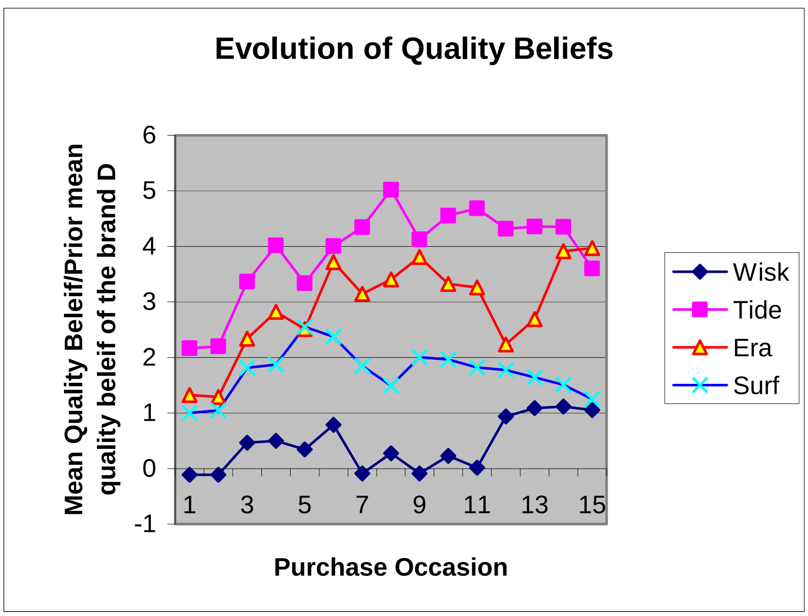
TABLE 8: Consideration Set Sizes when the respective Brands are chosen

(Ketchup Data Set)

Consideration Set Probabilities	Pr[# Brands = 1]	Pr[# Brands = 2]	Pr[# Brands = 3]	Pr[# Brands = 4]
Heinz chosen	0.362	0.269	0.251	0.118
Hunt's chosen	0.110	0.346	0.442	0.102
Del Monte chosen	0.181	0.270	0.376	0.173
Generic chosen	0.104	0.200	0.469	0.227

FIGURE1: Evolution of Mean Quality Beliefs for 15 Purchase Occasions

(Liquid Detergents)



REFERENCES

Andrews, R.L. and T.C. Srinivasan (1995), "Studying Consideration Effects in Empirical Choice Models Using Scanner Panel Data," *Journal of Marketing Research*, 32 (1), 30-41.

Allenby, G.M and J.L. Ginter (1995), "The effects of in-store Displays and Feature Advertising on Consideration Sets," *International Journal of Research in Marketing*, 12 (1), 67-81.

Bagwell, K. and M. Riordan (1991), "High and Declining Prices Signal Product Quality," *American Economic Review*, 81, 224-239.

Blattberg, R.C. and S.A. Neslin (1990), *Sales Promotions: Concepts, Methods and Strategies*, Prentice-Hall, New York.

Bronnenberg, B.J and W.R. Vanhonacker (1996), "Limited Choice Sets, Local Price Response, and Implied Measures of Price Competition," *Journal of Marketing Research*, 33 (2), 163-173.

Bucklin, R.E. and J.M. Lattin (1991), "A Two-State Model of Purchase Incidence and Brand Choice," *Marketing Science*, 10 (1), 24-39.

Chiang, J., S. Chib and C. Narasimhan (1999), "Markov Chain Monte Carlo and Models of Consideration Set and Parameter Heterogeneity," *Journal of Econometrics*, 89 (Mar/Apr), 223-248.

Chintagunta, P., D. Jain and N. Vilcassim (1991), "Investigating Heterogeneity in Brand Preference in Logit Model for Panel Data," *Journal of Marketing Research*, 28, 417-428**.**

Chintagunta, P.K. (1993), "Estimating a Multinomial Probit Model of Brand Choice Using the Method of Simulated Moments, *Marketing Science*, 11 (4), 1992, 386-407.

DeGroot, M.H. (1970), *Optimal Statistical Decisions*, McGraw-Hill, New York.

Desai, K.K. and W.D. Hoyer (2000), "Descriptive characteristics of memory-based consideration sets: Influence of usage Occasion frequency and usage location familiarity," *Journal of Consumer Research*, 27 (3), 309-323

Erdem, T. and M.P. Keane (1996), "Decision-making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," *Marketing Science*, 15 (1), 1-20.

Feinberg, F.M. and J. Huber (1996), "A Theory of Cutoff Formation under Imperfect Information," *Management Science*, 42 (1), 65-84.

Guadagni P. M. and J. D. C. Little (1983), "A Logit Model of Brand Choice Calibrated on Scanner Data," *Marketing Science*, 2 (3), 203-238.

Hauser, J.R. and B. Wernerfelt (1990), "An Evaluation Cost Model of Consideration Sets," *Journal of Consumer Research*, 16 (March), 393-408.

Horstmann, I. and G.M. MacDonald (1994), "When is Advertising a Signal of Product Quality?" *Journal of Economics and Management Strategy*, 3 (Fall), 561-584.

Howard, J.A. and J.N. Sheth (1969), *The Theory of Buyer Behavior*, New York: John Wiley.

Jeuland, A. P. (1979), "Brand Choice Inertia as one Aspect of the Notion of Brand Loyalty," *Management Science*, 25 (7), 671.

Johnson, N. and S. Kotz (1974), *Distributions in Statistics – Continuous Multivariate Distributions*, New York: John Wiley and Sons.

Kamakura, W.A. and G.J. Russell (1989), "A Probabilistic Choice Model of Market Segmentation and Elasticity Structure," *Journal Of Marketing Research*, 26 (4), 379-391.

Kotler, P. (1996), *Marketing Management: Analysis, Planning, Implementation and Control*, 9th Edition, Prentice-Hall.

Manrai, A.K and R.L. Andrews (1998), "Two-stage Discrete Choice Models for Scanner Panel Data: An Assessment of Process and Assumptions," *European Journal of Operational Research*, 111 (20), 193-215.

McFadden, D. (1981), "Econometric Models of Probabilistic Choice," in *Structural Analysis of Discrete Data with Econometric Applications*, (Eds.) C. Manski and D. McFadden, Cambridge, MA: MIT Press.

McFadden, D. (1989), "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration," *Econometrica*, 57, 995-1026.

Milgrom, P. and J. Roberts (1986), "Prices and Advertising Signals of Product Quality," *Journal of Political Economy*, 94, 796-821.
Mitra, A. (1995), "Advertising and the Stability of Consideration Sets over Multiple Purchase Occasions," International Journal of Research in Marketing, 12 (1), 81-94.
Moorthy, K.S. (1988), "Product and Price Competition in a Duopoly," *Marketing Science*, 7(2), 141-168.
Moorthy, K.S. and J. Banks (1999), "A Model of Price Promotions with Consumer Search," *International Journal of Industrial Organization*, 17, 371-398.
Moorthy, K.S. and K. Srinivasan (1995), "Signaling Quality with a Money-Back Guarantee: The Role of Transaction Costs," *Marketing Science*, 14 (1), 442-466.
Murthi, B.P.S. and K. Srinivasan (1999), "Consumer's Extent of Evaluation on Brand Choice," *Journal of Business,* 72 (2), 229-256.
Narasimhan, C. (1988), "Competitive Promotional Strategies," *Journal of Business*, 61 (October), 427-449.
Nedungadi, P. (1990), "Recall and Consumer Consideration Sets: Influencing Choice without Altering Evaluation," *Journal of Consumer Research*, 17 (December), 263-276.
Nelson, P. (1974), "Advertising as Information," *Journal of Political Economy*, 81, 729-754.
Newman, J.W. and R. Staelin (1972), "Pre-purchase Information Seeking for New Cars and Major Household Appliances," *Journal of Marketing Research*, 9 (August), 249-257.
Pakes, A. and D. Pollard (1989), "Simulation and the Asymptotics of Optimization Estimators," Econometrica, 57 (5), 1027.
Raju, J.S., V. Srinivasan and R. Lal (1990), "The Effects of Brand Loyalty on Competitive Price Promotional Strategies," *Management Science*, 36 (March), 276-304.
Ratchford, B. (1982), "Cost-Benefit Models for Explaining Consumer Choice and Information Seeking Behavior," *Management Science*, 28 (2), 197-212.
Robert, J. and D.O. Stahl II (1993) , "Informative Price Advertising in a Sequential Search Model," *Econometrica*, 61 (3), 657-686.
Roberts, J. H. and J.M. Lattin (1991), "Development and Testing of a Model of Consideration Set Composition," *Journal of Marketing Research*, 28 (4), 429-440.
Roberts, J. H. and J. Urban (1988), "Modeling Multi-attribute Utility Risk and Belief Dynamics," *Management Science*, 34 (2), 167-185.
Roy, R., P.K. Chintagunta and S. Haldar (1996), "A Framework for Analyzing Habits, "Hand-of-Past," and Heterogeneity in Dynamic Brand Choice," *Marketing Science*, 15 (3), 280-299.
Seetharaman, P. B., A. Ainslie and P.K. Chintagunta (1999), "Investigating Household State Dependence effects across Categories," *Journal of Marketing Research* 36 (4), 488-500.
Siddarth, S, R.E. Bucklin and D.G. Morrison (1995), "Making the Cut: Modeling and Analyzing Choice Set Restrictions in Scanner Panel Data," *Journal of Marketing Research,* 32 (8), 255-266.
Singleton, K.J. (1985), "Testing Specifications of Economic Agents' Inter-Temporal Optimum Problems in the Presence of Alternative Models," *Journal of Econometrics*, 30, 391-413.
Stigler, G.J. (1961), "The Economics of Information," *Journal of Political Economy*, 69, 213-235.
Wernerfelt, B. (1991), "Brand Loyalty and Market Equilibrium," *Marketing Science*, 10 (3), 229-245.

APPENDIX

PROPOSITION 1: *Greater the time consumer consecutively stays in a state of inertia, greater is the probability that she would add more brands into her consideration set on the next purchase occasion.*

PROOF: Consider the case when consumer *i* purchases brand *k* in a state of inertia for the first *t* purchase occasions. In other words, there was only brand *k* in the optimal consideration of consumer *i* set for *t* consecutive purchase occasions. The difference in the expected benefits at time *t+1* for choosing only brand *k* and for choosing brand *k* along with other set of brands {*j*} is:

$$EB_{i,\{j\}+k,t} - EB_{ik,t} = \beta \log\left(1 + \frac{\sum_{h\in\{j\}} \exp\left(\bar{u}_{ih,t}/\beta\right)}{\exp\left(\theta\omega_{ik,t+1} - \bar{p}_k/\beta\right)}\right) - \sum_{h\in\{j\}} C_{ih,t} \tag{A.2}$$

The set {*j*} will be the consumer's optimal consideration set if $EB_{i(j\}+k,t} - EB_{ik,t} \geq 0$. The probability of this event happening would be:

$$\Pr\left(EB_{i(j\}+k,t} - EB_{ik,t} \geq 0\right) = \Pr\left[\theta\omega_{ik,t+1} \geq \overline{P_k} + \beta \exp\left(\frac{1}{\beta}\sum_{h\in\{j\}} C_{ih,t}\right) - \beta \sum_{h\in\{j\}} \exp\left(\bar{u}_{ih,t}/\beta\right)\right]$$

$$= 1 - \Phi\left(\frac{\gamma\,\bar{p}_k + \beta \exp\left(\frac{1}{\beta}\sum_{h\in\{j\}} C_{ih,t}\right) - \beta \sum_{h\in\{j\}} \exp\left(\bar{u}_{ih,t}/\beta\right)}{\sigma\,\omega_{ik,t+1}}\right) \tag{A.3}$$

From equation (A.3), it can be seen that the variance of $\omega_{ik,t}$ is an increasing function of *t*. Therefore, the probability of adding additional brands in the consideration set is also an increasing function of *t*. Also, if other brands in the set {*j*} are on display or are advertised, their search costs $\sum_{\{j\}} C_{ij,t}$ would decrease. Hence, as *t* increases, more brands will be added to the consideration set.

■

PROPOSITION 2: *Greater the value of the variance of the utility function in the consideration stage,* $\sigma^2_{u_{ij,t}} = \sigma^2_{p_j}$ *, greater is the size of the optimal consideration set.*

PROOF: Consider the case when the variance of the prices of all the brands is the same (that is, $\sigma^2_{p_j} = \sigma^2_p$ for all brands *j*). Define $\beta = \sqrt{6}\gamma\sigma_p/\pi$ and $\bar{u}_{ij,t} = \omega_{ij,t} - \gamma\bar{P}_j$. Also define the expected benefit for choosing any consideration set {*h*} as $EB_{\{h\}} = E\max\left[\{u_{ik,t}\}_{k\in\{h\}} - \sum_{k\in\{h\}} C_{ik,t}\right]$. Consider a given value of $\beta = \beta'$ for which $\{m\}_{it}$ = {*j*} is the optimal consideration set. We need to prove that if the value of β increases, more brands will be added to the optimal consideration set $\{m\}_{it}$ (or a superset of {*j*} will be the optimal consideration set). Let $\{\tilde{j}\}$ be such a superset of {*j*} that contains an additional brand *l* that was not an element of the set {*j*} (in other words, $\{\tilde{j}\} = \{j\} \cup l$ and $l \in \{j\}$). Consider the difference of the expected for considering the consideration sets *{j}* and $\{\tilde{j}\}$

$$
\begin{aligned}
EB_{\{\tilde{j}\}} - EB_{\{j\}} &= \beta \log\left(\exp\left(\bar{u}_{il,t} / \beta\right) + \sum_{k \in \{j\}} \exp\left(\bar{u}_{ik,t} / \beta\right) \right) - C_{il,t} - \beta \log\left(\sum_{k \in \{j\}} \exp\left(\bar{u}_{ik,t} / \beta\right) \right) \\
&= \beta \ln\left(1 + \frac{\exp\left(\bar{u}_{il,t} / \beta\right)}{\sum_{k \in \{j\}} \exp\left(\bar{u}_{ik,t} / \beta\right)} \right) - C_{il,t} \qquad \text{(A.1)}
\end{aligned}
$$

Since $\{j\}$ was the optimal consideration set for $\beta = \beta'$, $EB_{\{\tilde{j}\}} - EB_{\{j\}} \leq 0$. Thus, $EB_{\{\tilde{j}\}} - EB_{\{j\}}$ is monotonically increasing in β and $EB_{\{\tilde{j}\}} - EB_{\{j\}} \to \infty$ when $\beta \to \infty$. Since $EB_{\{\tilde{j}\}} - EB_{\{j\}}$ is a continuously differentiable function of β, there exists a value $\beta > \beta'$, when $EB_{\{\tilde{j}\}} - EB_{\{j\}} > 0$. Therefore, as the value of β increases, the size of the optimal consideration set will increase.

■