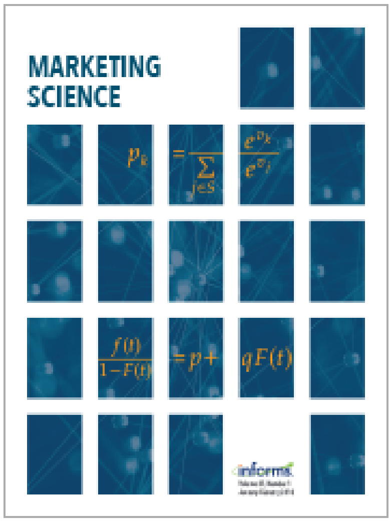



Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

# Price Uncertainty and Consumer Search: A Structural Model of Consideration Set Formation

Nitin Mehta, Surendra Rajiv, Kannan Srinivasan,

# Price Uncertainty and Consumer Search: A Structural Model of Consideration Set Formation

Nitin Mehta • Surendra Rajiv • Kannan Srinivasan
*Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, Ontario, Canada M5S 3E6*
*NUS Business School, National University of Singapore, BIZ 1, 1 Business Link, Singapore 117592*
*Graduate School of Industrial Administration, Carnegie Mellon University,*
*5000 Forbes Avenue, Pittsburgh, Pennsylvania 15213*
*nmehta@rotman.utoronto.edu.ca • bizar@nus.edu.sg • kannans@andrew.cmu.edu*

We offer an econometric framework that models consumer's consideration set formation as an outcome of her costly information search behavior. Because frequently purchased products are characterized by frequent price promotions of varying depths of discounts, a consumer faces significant uncertainty about the prices of the brands. The consumers engage in a fixed-sample search strategy that results in their discovering the posted prices of a subset of the available brands. This subset is referred to as the consumer's "consideration set."

The proposed model is estimated using the scanner data set for liquid detergents. Our key empirical results are: (i) consumers zincur significant search costs to discover the posted prices of the brands; (ii) whereas in-store displays and feature ads do not influence consumers' quality perceptions of the brands, they significantly reduce search costs for observing the prices of the brands; (iii) per capita income of consumer's household significantly increases her search costs; and (iv) the consumers' price sensitivity is seriously underestimated if we were to assume that consumers get to know all the posted prices at zero cost. The proposed model is also estimated for the ketchup category to enable us to do cross-category comparisons of consumers' price search behavior.

(*Price Uncertainty; Consumer Search; Consideration Set; Quality Uncertainty; Consumer Learning; Bayesian Updating; Structural Model; Econometric Estimation*)

## 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 (because 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



0732-2399/03/0022/0058$5.00
1526-548X electronic ISSN

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 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. Because 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 to get a more precise estimate of the associated consumer surplus.
- 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 §1.2, typically assume that consumers use 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; furthermore, consumer brand choice decision is arrived at by explicit utility comparison across all the brands in the product category. In addition, because 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 with 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 because these products are more "experience or credence goods" (Kotler 1996). As such, consumers cannot resolve quality uncertainty by engaging in explicit prepurchase search. Because the FPPs are frequently promoted, consumers do not have complete information about their price: Although 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, although they are aware of the price distributions. Consumer rationality implies that consumers will engage in price search to reduce this uncertainty. However, because 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."*[1] The consideration set formation and the final choices are modeled probabilistically. The optimal consideration set and search costs are endogenously estimated from a scanner panel data set. Also, consumers are uncertain about the 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 approaches proposed by Jovanovic (1979), Miller (1984), and Erdem and Keane (1996).

Our structural model yields the following insights about the size of the optimal consideration set and consumers' search behavior in FPP categories:

(i) Relative to consumers with low-price sensitivity, consumers with high-price sensitivity 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.

We calibrate the proposed model using the ERIM data set for the liquid detergent product category. To analyze cross-category variations in consumer price search behavior, we also calibrate the model using the 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, the estimated baseline search cost is $C_0 = 0.0472$, and the average consideration set size is 2.31.

(ii) For the ketchup category, the estimated baseline search cost is $C_0 = 0.0284$, and the average consideration set size is 2.28. The reason that the consideration set sizes are the same in the two categories despite lower search cost for ketchup is that price variation in the ketchup category is less than price variation in the liquid detergent category. As a result, the expected benefit of search in liquid detergents is more than in the ketchup category. This factor offsets the lower baseline search costs in the ketchup category.

(iii) Although in-store display activities and feature ads do not influence quality perceptions, they do reduce consumer search costs for a brand, thereby significantly increasing the probability of the brand being considered. We find feature ads to be more effective than displays in influencing consumer search behavior for both liquid detergents and ketchup. Comparing across the two categories, we find feature ads to be relatively more effective in influencing search behavior for liquid detergents. In contrast, we find the impact of displays on the intensity of consumer search to be more for ketchup than for liquid detergents.

(iv) We find that, for the liquid detergent data set, per capita income significantly increases consumers' search costs. We also find that consumers tend to engage in more extensive price search during weekends than during weekdays.

(v) Prior consumption experiences influence quality perceptions of the brands for both the liquid detergents and the ketchup categories. However, consumption experiences yield only limited additional quality information. This is consistent with our expectation, because these categories are mature.

(vi) The consumers' price sensitivity is seriously underestimated if we were to assume that consumers get to know the posted prices of all the brands at a zero cost.

[1] We recognize that the extant literature mostly deals with out-of-store behavior in which consideration sets are linked to brand attributes. In particular, Gensch (1987) posits that consumers may use disjunctive rules ("elimination-by-aspect" heuristics) to limit the number of brands that they evaluate in detail. We agree it is possible that consumers follow more complex screening schemes. At the first stage, they select the subset of brands that meet the minimal criteria on a set of attributes. Let us denote this set as the "acceptance set." It is clear that, on a shopping trip, a consumer does not have any incentive to search for posted prices of brands outside this "acceptance set" (after all, why would I look for price discounts on a brand that I know I would never buy even if it was available for free?). However, if the "acceptance set" is large, a consumer might not consider all the brands in this set on all purchase occasions. Thus, at the second stage, the consumer decides on the "consideration set" based on information search considerations. Thus, we see the attribute-based "elimination-by-aspect" models and the proposed search-based consideration set model as complementary.

(vii) We find that models that ignore differences in consumer's intrinsic preference for the various brands seriously overestimate the consideration set sizes. Unlike the proposed model, extant models of consideration set (e.g., Andrews and Srinivasan 1995, Siddarth et al. 1995, and Bronnenberg and Vanhonacker 1996) do not include the brand intercept term in the utilities in the consideration stage.

### 1.2. Related Literature and Research Contributions

Some recent studies have attempted to incorporate the notion of consideration or choice sets in the context of stochastic choice models. In a pioneering research effort, Hauser and Wernerfelt (1990) propose a theoretical evaluation cost model of consideration sets. The basic idea here is that consumers face uncertainty about the utilities associated with the various brands and, furthermore, that by engaging in a costly preevaluation 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 on the brands' price range membership. Siddarth et al. (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 et al. (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 really means and why "consideration utility" should be different from "choice utility." Similarly, the specification from Chiang et al. (1999) takes a purely combinatorial approach to consideration sets and is not motivated by any behavioral assumption. Although this framework allows for consumer heterogeneity in consideration sets, the consideration set for an individual consumer 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 the notion of limited consumer search in a reduced-form approach. Bucklin and Lattin (1991) propose a two-state model of consumer brand choice that explicitly models two types of shopping behavior, namely, planned and unplanned purchases. They posit that, 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-purchase interventions, such as price promotions. 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 (although 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. 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 remainder of the paper is arranged as follows. In §2, we develop the model formulation. Section 3 describes the data set used for calibrating the proposed model, discusses the estimation method, reports the parameter estimates, and discusses the managerial implications of our findings. Section 4 concludes.

## 2. Model Development

In this section, we lay out the model specification and discuss the main assumptions underlying the proposed formulation. Before developing the econometric specification, in §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 to derive an identifiable statistical model of consideration set formation. Finally, §2.5 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) variability of prices.

[2] Analytical details are more elaborately discussed in the Technical Appendix that is available from http://mktsci.pubs@informs.org.

### 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 of uncertainty is captured by hypothesizing that the consumer has prior beliefs about the 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. Consumption experience does not fully reveal the intrinsic brand qualities. Thus, 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 whereas 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) Because the brands in the product category are occasionally on "sale," and because the depth of discount offered during price promotions varies across sale events, a consumer is uncertain about the posted price of the brands on any purchase occasion. However, we assume that through prior 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,

[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.

*"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 whereas 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 and 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 $\theta$ denotes the consumer's intensity of preference for quality (alternatively, her marginal willingness to pay for quality). The above specification for the indirect utility can be derived from a direct utility specification called the "cross-product repackaging utility."[4] Such a specification has been used by Blattberg and Wisniewski (1989) to explain asymmetric switching between brands and has been traditionally employed in the vertical differentiation literature (e.g., Moorthy 1988, Bagwell and Riordan 1991, Kalra et al. 1998, Villas-Boas 1998).

We assume further 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* that are captured by her priors $f(q_{ijt})$. In particular, we assume a normal distribution for consumer prior beliefs, i.e.,

$$q_{ij,t} \mid \mathsf{H}_i(t) \sim N\big(\omega_{ij,t}, \sigma^2_{\omega_{ijt}}\big), \quad \forall i, \forall j. \tag{2}$$

Note that $\omega_{ij,t}$ denotes consumer $i$'s current estimate of the expected quality of brand $j$ given her consumption history, $\mathsf{H}_i(t)$. $\sigma^2_{\omega_{ijt}}$ 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. This belief is assumed to be identical across all consumers. Thus, we assume that initial quality beliefs are given by

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

We assume that a consumer learns about the intrinsic quality of a brand through consumption experiences. We assume further 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. Third, the consumers might even forget about the qualities of the brands they had learned during the previous consumption experiences.[6] The net impact of this assumption is that whereas a consumer's uncertainty

[4] Please see R. D. Willig (1978) and W. M. Hanemann (1984) for additional details.

[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] We thank an anonymous reviewer for suggesting this as a possible reason for quality 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).[7]

Let $q_{ij,t-1} \sim N(\omega_{ij,t-1}, \sigma^2_{\omega_{ij,t-1}})$ 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

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

In Equation (4), $q_j$ denotes the "true" quality of brand $j$, whereas the random variable $\eta_{ij,t-1}$ denotes the noise associated with consumption experience with $\eta_{ij,t-1} \sim N(0, \sigma^2_\eta)$.[8] Thus, by consuming brand $j$ in period $t-1$, the consumer gets to learn more about the "true" quality of the brand. However, because the signal is noisy, her expectation about brand $j$'s quality, $\omega_{ij,t-1}$, does not get updated to $q_j$. Note that $\sigma^2_\eta$ is a measure of the informativeness of consumption experience: $\sigma^2_\eta = 0$ 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, $\sigma^2_\eta$, to be the same across all the brands and over time.

We assume that the consumer uses the realized value of the consumption signal, $\hat{\lambda}_{ij,t-1}$, to update her prior beliefs about brand $j$'s quality, $q_{ij,t-1}$, in a Bayesian manner. 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), the mean and variance of the posterior beliefs are related to the mean and variance of the prior beliefs as follows:

$$\begin{aligned} \omega_{ij,t} &= \frac{\omega_{ij,t-1}\alpha_{ij,t-1} + d_{ij,t-1}\hat{\lambda}_{ij,t-1}}{\alpha_{ij,t-1} + d_{ij,t-1}}; \quad \text{and} \\ \alpha_{ij,t} &= \alpha_{ij,t-1} + d_{ij,t-1}. \end{aligned} \tag{5}$$

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

[8] 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.

In Equation (5), we use $\alpha_{ij,t-1} = \sigma^2_\eta / \sigma^2_{\omega_{ij,t-1}}$.[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 $j$.

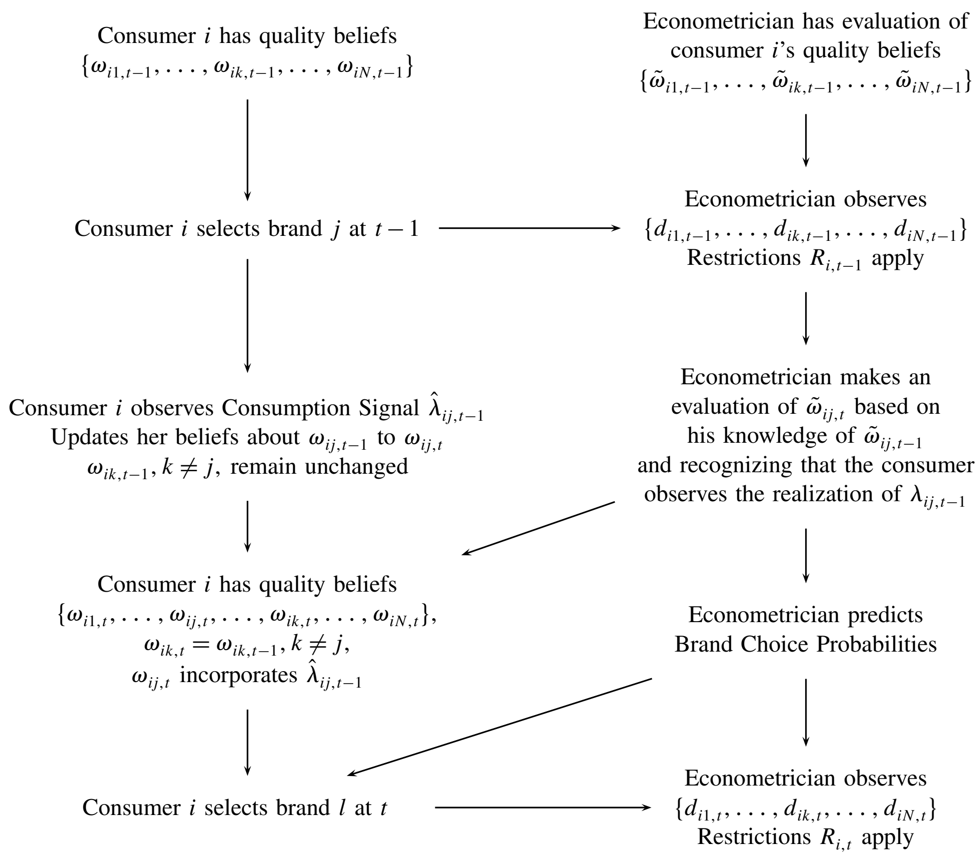
Equation (5) characterizes the law of motion of the mean of consumer $i$'s subjective quality beliefs about brand $j$ as she receives the consumption signals of brand $j$. As shown in the Technical Appendix, the mean of the consumer's quality beliefs converges to the true quality $q_j$; i.e., $\omega_{ij,t} \to q_j$ as $t \to \infty$. It is important to note that although the consumer observes the realizations of consumption signals $\{\hat{\lambda}_{ij,s}\}_{s=1}^{t-1}$ (and hence knows deterministically the values of $\omega_{ij,t}$ for all brands $j$ at any time $t$), the analyst does not. In the absence of observations on $\{\hat{\lambda}_{ij,s}\}_{s=1}^{t-1}$, from the analyst's perspective, $\omega_{ij,t}$ is a normal random variable that we denote by $\tilde{\omega}_{ij,t}$. Hence, the analyst can only make probabilistic statements about the consumer's consideration set and brand choice decisions.[10]

To help better understand the asymmetry in the information sets of the consumer and the analyst, we graphically represent the sequence of events in Figure 1. As shown in the Technical Appendix, from the analyst's perspective the signal $\lambda_{ij,t-1}$ is distributed as $\lambda_{ij,t-1} \sim N(\tilde{\omega}_{ij,t-1}, \sigma^2_\eta + \alpha^{-1}_{ij,t-1})$. Further, because we set $\sigma^2_\eta = 1$ for identification purposes, the signal $\lambda_{ij,t-1}$ is distributed as $\lambda_{ij,t-1} \sim N(\tilde{\omega}_{ij,t-1}, 1 + \alpha^{-1}_{ij,t-1})$. Substituting $\lambda_{ij,t-1} \sim N(\tilde{\omega}_{ij,t-1}, 1 + \alpha^{-1}_{ij,t-1})$ in Equation (5),

[9] Because we can only estimate $\alpha_{ij,t-1} = \sigma^2_\eta / \sigma^2_{\omega_{ij,t-1}}$, for identification purposes, we set the variance of consumption signal to be one; i.e., $\sigma^2_\eta = 1$.

[10] This is similar to the case of discrete choice models (McFadden 1981). The consumer knows her utility for the different brands and hence can deterministically determine which brand yields the highest utility. In contrast, the analyst does not know the random components of the consumer's utility function and hence can only make probabilistic statements about consumer brand choice.

**Figure 1 Sequence of Events and Information Sets**




we get the analyst's estimate of $\tilde{\omega}_{ij,t}$ conditional on $\tilde{\omega}_{ij,t-1}$ as

$$\tilde{\omega}_{ij,t} = \tilde{\omega}_{ij,t-1} + N\left(0, \frac{d_{ij,t-1}(1+\alpha_{ij,t-1}^{-1})}{(\alpha_{ij,t-1} + d_{ij,t-1})^2}\right). \tag{6}$$

Note that in Equation (6), the first term $\tilde{\omega}_{ij,t-1}$ is a truncated normal random variable, subject to restrictions imposed in the error space because of consumer $i$'s brand choices in periods $s = 1$ till $t-1$. These restrictions on the error space are discussed in detail in the Technical Appendix. Also, note that the second term in Equation (6) 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 impacts on quality perception. 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 et al. 1999) updating of quality perceptions over successive consumption experiences. Note that the Guadagni and Little (1983) loyalty term allows for the loyalty to a brand to fall if the brand is not consumed. But, if the brand is consumed, then the loyalty of the brand will always increase (or always decrease depending on the sign of the loyalty coefficient). On the other hand, in the Bayesian updating model, the preference of the brand can increase, decrease, or stay the same when the brand is purchased (the increase or decrease depends on the magnitude of the noisy quality signal received from consumption).

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 (6), the analyst's evaluation of consumer $i$'s mean quality perception of brand $j$ given her purchase history $\mathsf{H}_i(t)$, is given by $\tilde{\omega}_{ij,t} = N_{ij,t-1} + \sum_{l=0}^{t-2}$ Truncated $N_{ij,l}$. We reiterate 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, because she gets to observe these realized values of the consumption signals, namely, $\lambda_{ij,t-1}$, $\forall t$. However, from the perspective of the analyst, $\tilde{\omega}_{ij,t}$ 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 (6) is that the random error component exhibits serial correlation, i.e., "habit persistence" (Roy et al. 1996). Furthermore, because the variance of $\tilde{\omega}_{ij,t}$ 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 raises questions about the standard assumptions of homoskedasticity and absence of serial correlation routinely made in the literature. Note 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, $\mathsf{H}_i(t)$, her updated (posterior) beliefs about qualities of the competing brands are denoted by $q_{ij,t} \sim N(\omega_{ij,t}, \sigma^2_{\omega_{ij,t}})$, $\forall j$. 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 because retailers routinely offer price promotions of varying depths of discount, the actual price prevailing on any shopping trip is a random variable. Although the consumer does not know the actual posted price, we assume that she is aware of price distributions for the various brands denoted by $p_{ijt} \sim f_j(\bar{p}_j, \sigma^2_{p_j})$, $\forall j$, 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 (8) later.

Given the twofold 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 $\theta\omega_{ij,t} - \bar{p}_{ij,t}$. 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 $\theta\omega_{ij,t} - \bar{p}_{ij,t}$ and variance $\sigma^2_{uj} = \sigma^2_{pj}$; i.e.,

$$U_{ij,t} \sim EV\left(\frac{\pi}{\sqrt{6}\sigma_{uj}}, \frac{\sqrt{6}\sigma_u e_c}{\pi} - \theta\omega_{ij,t} + \bar{p}_j\right), \qquad (7)$$

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

Although 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 was 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

[11] Of course, subject to the restrictions imposed by the choices made by consumer $i$ over her consumption history (i.e., from $t = 0$ to $t-1$). This gives the analyst the information about relative magnitudes of $\tilde{\omega}_{ij,t}$ for all the $j$ brands. Note that $\tilde{\omega}_{ij,t} \to q_j$ as $t \to \infty$, because the number of identifying restrictions become infinite.

[12] Specifically, adding an i.i.d. error term with zero mean and finite variance would make the model unidentifiable. Although 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.

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 because of 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 decide 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, $\mathsf{H}_i(t)$. 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 yield the highest net expected surplus (net of the total search cost).

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, because 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). 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 $\sigma^2_{p_j}$ is the 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 (7) is the same across all the brands and remains constant over a consumer's purchase history; i.e., $\sigma^2_{ujt} = \sigma^2_u = \sigma^2_p$, $\forall i, \forall j, \forall t$.[14]

Consider, hypothetically, the consumer selecting brands 1 to $N$ with $N \leq J$ 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]

$$EB^{\{j\}}_{i,t} = \frac{\sqrt{6}\sigma_u}{\pi} \ln\left(\sum_{l=1}^{N} \exp\left(\frac{\pi}{\sqrt{6}\sigma_u}(\theta\omega_{il,t} - \bar{p}_j)\right)\right) - \sum_{l=1}^{N} C_{il,t}. \tag{8}$$

Note that the expected benefit in equation (8) 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

$$\{k\} = \underset{\{j\}}{\operatorname{argmax}} \; \frac{\sqrt{6}\sigma_u}{\pi} \ln\left(\sum_{l\in\{j\}} \exp\left(\frac{\pi}{\sqrt{6}\sigma_u}(\theta\omega_{il,t} - \bar{p}_l)\right)\right) - \sum_{l\in\{j\}} C_{il,t}. \tag{9}$$

In Equation (9), because the expected qualities $\omega_{il,t}$ is known to the consumer, she can determine which of the sets $\{j\}$ yields the highest expected benefit. However, from the analyst's perspective $\tilde{\omega}_{ij,t}$ are random variables (because the analyst does not observe the

[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 will not 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 four brands (see Table 1a). In fact, as pointed out by the area editor, we can demonstrate this using a simple $F$-test on the variances of the different brands. As a result, we have used Equation (8) 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.

realized values of the quality cues). As such, given the consumer's consumption history at time $t$, $\mathsf{H}_i(t)$, we can only write the probability that the set $\{k\}$ is the optimal consideration set:

$$\begin{aligned}
&\Pr[\{k\}=\text{Consideration Set}]\\
&=\Pr\Bigg[\frac{\sqrt{6}\sigma_u}{\pi}\ln\Bigg(\sum_{l\in\{k\}}\exp\Bigg(\frac{\pi}{\sqrt{6}\sigma_u}\left(\theta\tilde{\omega}_{il,t}-\bar{p}_l\right)\Bigg)\Bigg)-\sum_{l\in\{k\}}C_{il,t}\\
&\geq\frac{\sqrt{6}\sigma_u}{\pi}\ln\Bigg(\sum_{l\in\{j\}}\exp\Bigg(\frac{\pi}{\sqrt{6}\sigma_u}\left(\theta\tilde{\omega}_{il,t}-\bar{p}_l\right)\Bigg)\Bigg)-\sum_{l\in\{j\}}C_{il,t}\Bigg],\\
&\forall\{j\}. \qquad (10)
\end{aligned}$$

**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, prior to the consumer making a brand choice, the price uncertainty gets completely resolved for all the brands under consideration. However, she still 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, because $\tilde{\omega}_{ij,t}$s are random variables from the analyst's perspective, the analyst can only make a probabilistic statement about the optimal chosen brand. In particular, because $\tilde{\omega}_{ij,t}$s are truncated normal random variables [compare Equation (6)],[16] the probability that the consumer selects brand $j$, given that $j\in\{k\}$, is given by a Probit probability:

$$\begin{aligned}
&\Pr\left(\delta_{ij,t}=1\mid j\in\{k\}\right)\\
&\quad=\Pr\left(\theta\tilde{\omega}_{ij,t}-p_{ij,t}\geq\theta\tilde{\omega}_{ih,t}-p_{ih,t},\forall h\in\{k\}\mid \mathsf{H}_i(t)\right)\\
&\quad=\int_{-\infty}^{p^1_{ij,t}/\theta}\cdots\int_{-\infty}^{p^N_{ij,t}/\theta}\text{Truncated }\phi\big(\tilde{\omega}^1_{ij,t},\cdots,\tilde{\omega}^{j-1}_{ij,t},\\
&\qquad\qquad\tilde{\omega}^{j+1}_{ij,t},\cdots,\tilde{\omega}^{N}_{ij,t}\big)\\
&\qquad\times\prod_{\substack{h\in\{k\}\\ h\neq j}}d\tilde{\omega}^h_{ij,t}. \qquad (11)
\end{aligned}$$

[16] In addition to the restrictions imposed because of the consumer's sequence of past choices, the variables $\tilde{\omega}_{ij,t}$ also need to satisfy restrictions because brand $j$ belongs to the optimal consideration set $\{k\}$. Please see the Technical Appendix for additional details. We thank the area editor for reminding us to point this out.

In Equation (11), 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 $\tilde{\omega}^h_{ij,t}$ and $p^h_{ij,t}$ as follows: $\tilde{\omega}^h_{ij,t}=\tilde{\omega}_{ih,t}-\tilde{\omega}_{ij,t}$ 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:

$$\begin{aligned}
\Pr\left(\delta_{ij,t}=1\right)=\sum_{\{k\}}\Pr\left(\delta_{ij,t}=1\mid j\in\{k\}\right)\\
\times\Pr\left[\{k\}=\text{Consideration Set}\right]. \qquad (12)
\end{aligned}$$

### 2.3. Controlling for Unobserved Consumer Heterogeneity

In the proposed formulation, because 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 et al. 1991, Gonul and Srinivasan 1993) as it allows for:

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

We are able to accomplish this finer variance decomposition, because $\omega_{ijt}$ needs to satisfy certain conditions over the consumption history $\mathsf{H}_i(t)$ that are discussed in detail in the Technical Appendix.

To account for heterogeneity in the consumer's intensity of preference (marginal willingness to pay) for quality, we assume $\theta$ to be gamma distributed across the consumer population with mean $\overline{\theta}$ and variance $\sigma^2_\theta$. This ensures that the consumer's marginal willingness to pay for quality is positive, consistent with the vertical differentiation literature.[17]

[17] This is equivalent to allowing consumers to differ in their price sensitivity. Furthermore, assuming $\theta$ to be gamma distributed ensures that negative price sensitivity for all consumers.

### 2.4. Quality Uncertainty and Model Identification

As noted earlier, the randomness in $\omega_{ij,t}$ (from the analyst's perspective) is the source of randomness in our formulation and is akin to the random utility component in standard discrete choice models. If we were to use a "statistical error term" (e.g., as in Lattin and Roberts 1991) instead of quality learning to build our statistical model, then we would not be able to identify and estimate all the parameters at the consideration stage in our framework. Specifically, we would not be able to identify separately the baseline 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 was 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

$$\begin{aligned}
&\Pr[\{k\} = \text{Cons. Set}] \\
&\quad = \Pr\Bigg[\{k\} = \arg\max_{\{j\}} \frac{\sqrt{6}\sigma_u}{\pi} \\
&\qquad \times \ln\left(\sum_{l\in\{j\}} \exp\left(\frac{\pi}{\sqrt{6}\sigma_u}\left(\theta q_l - \bar{p}_l + \varepsilon_{il,t}\right)\right)\right) \\
&\qquad - \sum_{l\in\{k\}} C_{il,t}\Bigg].
\end{aligned}$$

This can be rewritten as

$$\begin{aligned}
&\Pr[\{k\} = \text{Cons. Set}] \\
&\quad = \Pr\Bigg[\{k\} = \arg\max_{\{j\}} \sum_{l\in\{j\}} \exp\Bigg(\frac{\pi}{\sqrt{6}\sigma_u} \\
&\qquad \times \left(\theta q_l - \bar{p}_l - \sum_{l\in\{j\}} C_{il,t} + \varepsilon_{il,t}\right)\Bigg)\Bigg]. \qquad (13)
\end{aligned}$$

We can see that there are two intercept parameters in the probability formulation in Equation (13): 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 a 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 §2.2) instead of adding the statistical error term, the probability of consideration in this case would be:

$$\begin{aligned}
&\Pr[\{k\} = \text{Cons. Set}] \\
&\quad = \Pr\Bigg[\{k\} = \arg\max_{\{j\}} \sum_{l\in\{j\}} \exp\Bigg(\frac{\pi}{\sqrt{6}\sigma_u} \\
&\qquad \times \left(\theta \tilde{\omega}_{il,t} - \bar{p}_l - \sum_{l\in\{j\}} C_{il,t}\right)\Bigg)\Bigg]. \qquad (14)
\end{aligned}$$

Recall from Equation (6) that the mean perceived quality of brand $l$ for consumer $i$ on purchase occasion $t$, $\tilde{\omega}_{il,t}$, in Equation (14) is the sum of $t$ normal random variables: $\tilde{\omega}_{ij,t} = N_{ij,t-1} + \sum_{l=0}^{t-2} \text{Truncated } N_{ij,l}$. Because the variance of the distribution of $\tilde{\omega}_{il,t}$ evolves over time from consumption experiences, we can identify separately the parameters $\alpha_0$ in the variance of $\tilde{\omega}_{il,t}$ and the search cost $C_{il,t}$.

### 2.5. 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 1. *The higher the variance of* (*indirect*) *utility, the larger is the size of the optimal consideration set.*

An immediate implication of Proposition 1 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 will exhibit higher search propensity. As such, the intensity of competition in these markets will be higher. We summarize this discussion in the following corollary.

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 with 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 1 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 a greater reliance on price promotions by firms induces more active price search by the consumers. That, in turn, would mean more intense price competition in the future. We summarize this intuition in the following corollary.

Corollary 2. *The greater the price variability in a product category, the 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 four 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 observations 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 other 200 households form the holdout sample to test the predictive validity of the proposed and competing specifications. The estimation sample had a total of 3,592 purchase observations, and the holdout sample had a total of 3,019 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.

**Table 1a** **Descriptive Statistics for the 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.

In addition, we do a cross-category analysis using the ketchup data set.[18] For our analysis, we have taken four brands: Heinz, Hunt's, Del Monte®, and the generic brand. The data set is comprised of 150 randomly selected households with a total of 2,332 observations. The minimum number of purchase observations 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. Similar to the liquid detergent category, we picked out 100 random households with a total of 1,532 observations to test the predictive validity of the proposed and competing specifications. The summary statistics for the data set are given in Table 1b.

In the preestimation 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

[18] The liquid detergent data set has been previously used by Siddarth et al. (1995). Bronnenberg and Vanhonacker (1996) use a European data set for dry detergents. The ketchup data set has been used earlier by Andrews and Srinivasan (1995) and Chiang et al. (1999).

**Table 1b Descriptive Statistics for the 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) |
| Hunt's | 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.

quality beliefs to be the same across all brands and across all consumers and without loss of generality set this belief to zero, i.e., $\omega_{ij,0} = 0\ \forall i, \forall j$. We then calibrate the model for the preestimation 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 the liquid detergent data set are given in Table TA.1 of the Technical Appendix and for the ketchup data set in Table TA.2 of the Technical Appendix.

**Factors Affecting Cost of Price Search.**[19] 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 greater the time spent to do so, the higher will be the consumer's search cost. Similarly, the higher the consumer's opportunity cost of time, the 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 $t$ depend 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.

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

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

In Equation (15), $DISPLAY_{ij,t}$ = 1 if brand $j$ is displayed on purchase occasion $t$ and = 0 otherwise. The indicator variable $FEATURE_{ij,t}$ is similarly defined. Because displays and feature ads give price information (Blattberg and Neslin 1990), they automatically reduce price search costs to zero.[20] However, we would not expect every consumer to observe displays and feature ads. Because we do not have the secondary data regarding the information a given individual shopper would acquire from a feature ad or a display, we must model each consumer as being exposed to the stimuli. 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) in the presence of feature ads or displays.

The time spent by a consumer to ascertain the posted price of a brand would also depend on the consumer's familiarity with its shelf location. In Equation (15), we capture the consumer's familiarity with the shelf location of the brands through two variables, namely, $STLOY_i$ 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,

[19] We thank the area editor and two anonymous reviewers for several excellent suggestions on modeling consumer search costs.

[20] In the liquid detergent category, brands are often feature advertised and displayed. The frequency of feature advertisements varies from 7% to 14% across the four selected brands. Similarly, the frequency of displays varies from 6% to 22% across the brands.

the less time consuming is the price search process (Desai and Hoyer 2000). Following Murthi and Srinivasan (1999), store and category familiarity 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. $DISPLAY_{ij,t-1}$ is a dummy variable that indicates whether brand $j$ was on display on previous purchase occasion $t$-1. The argument here is that because displays are usually placed in special locations, if on the previous purchase occasion the brand was on display, search for the brand will be more costly if it is not again displayed in the current period.

We capture the impact of opportunity cost of time on search costs through three variables, namely, $DOW_{it}$, $FTHM_i$, 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 is zero if the purchase was made during a weekday. It captures the shopping time available to the consumer. On weekends, because 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 $FTHM_i = 0$ 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 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:[21] $C_0 > 0$; $C_1 < 0$; $C_2 < 0$; $C_3 < 0$; $C_4 < 0$; $C_5 < 0$; $C_6 > 0$; and $C_7 > 0$.

[21] 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.

### 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, $\sigma_\theta^2$. We assume quality sensitivity parameter $\theta$ to be gamma distributed with mean $\bar{\theta}$ and variance $\sigma_\theta^2$;

(iii) Ratio of the noise in consumption signal to the information that can be gained at the beginning of the observation period, $\alpha_0$;

(iv) The search cost parameters, namely,

(a) $C_0$—Baseline search cost per brand for a household with per capita income of \$10,000 in the absence of feature advertisements 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 in the consumer's household on her 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.

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

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

**Table 2 Parameter Estimates for Models I and II for the Liquid Detergent Data Set**

| Parameter | Explanation | MODEL I (Std. Dev.) | MODEL II (Std. Dev.) |
|---|---|---|---|
| $\overline{\theta}$ | Mean value of the intensity of preference (willingness to pay) for quality across the consumers | 2.4168 (0.130) | 4.7522 (0.141) |
| $\sigma_\theta^2$ | Variance of the intensity of preference (willingness to pay) for quality across the consumers | 0.4251 (0.271) | 1.2199 (0.254) |
| $\kappa_f$ | Ratio of the informativeness (about the true quality) of feature advertisement signal to that of consumption signal | | 0.062 (0.151) |
| $\kappa_d$ | Ratio of the informativeness (about the true quality) of displays to that of consumption signal | | 0.041 (0.010) |
| $\alpha_0$ | Inverse of the uncertainty in quality of a brand 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) | |
| $C_4$ | 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 $1,000 on search costs | 0.0018 (0.0007) | |
| $C_7$ | Effect of display on previous purchase occasion on search costs | 0.0009 (0.0012) | |
| $\overline{\omega}_{\text{Wisk}}$ | Mean value of perceived quality of Wisk (across all consumers and across all purchase occasions) | 0.0097 (0.006) | 0.0431 (0.007) |
| $\overline{\omega}_{\text{Tide}}$ | Mean value of perceived quality of Tide (across all consumers and across all purchase occasions) | 0.1729 (0.008) | 0.0624 (0.008) |
| $\overline{\omega}_{\text{Era}}$ | Mean value of perceived quality of (across all consumers and across all purchase occasions) | 0.1135 (0.015) | 0.0469 (0.014) |
| $\overline{\omega}_{\text{Surf}}$ | Mean value of perceived quality of Surf (across all consumers and across all purchase occasions) | 0.0589 (0.003) | 0.0307 (0.003) |

(i) Model II is a variant of the proposed model (Model I) 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.[22] The parameters that we estimate in Model II are:

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

(b) Variance of quality sensitivity across the population, $\sigma_\theta^2$;

(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, $\alpha_0$;

(d) Ratio of the noise in consumption signal to the noise in feature advertisement, $\kappa_f$;

(e) Ratio of the noise in consumption signal to the noise in displays, $\kappa_d$.

The parameter estimates for Model II for the liquid detergents are given in Table 2 (Column 3):

(ii) Model AS is the model proposed by Andrews and Srinivasan (1995) incorporates 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. This model will be referred to as AS (1995) henceforth. The parameter estimates for AS (1995) for the liquid detergent data set are summarized in Table TA.3 of the Technical Appendix.

(iii) Model SBM is the model proposed by Siddarth et al. (1995) incorporates 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

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

of price, display, feature ads, and brand loyalty. This model will be referred to as SBM (1995) henceforth. The parameter estimates for the model for the liquid detergent data set are summarized in Table TA.4 of the Technical Appendix.

(iv) Model BV is the model proposed by Bronnenberg and Vanhonacker (1996) incorporates consideration and choice stages in a reduced-form framework. At the consideration stage, only the salience features of a brand—namely, 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). This model will be referred to as BV (1996) henceforth. The parameter estimates for the model for the liquid detergent data set are summarized in Table TA.5 of the Technical Appendix.

We do not compare the proposed specification with the framework developed in Lattin and Roberts (1991) because we would need self-reported consumer data to estimate their model. We do not compare the proposed model with Chiang et al. (1999) because it requires hierarchical Bayesian estimation methods while we use the MSM estimation approach.

Table 3a summarizes the results from comparison of the proposed model (Model I) against the four competing models (Model II, AS 1995, SBM 1995 and BV 1996) for the liquid detergent data set. We use the statistical test proposed by Singleton (1985) that is appropriate for comparing nonnested specifications in the Generalized Method of Moments framework.[23] 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 sample ($p$-value $< 0.02$) and the holdout sample ($p$-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 the models AS (1995) and SBM (1995) for both the estimation sample ($p$-value $< 0.1$) and the holdout sample ($p$-value $< 0.05$). This allows us to infer that, relative to the competing models AS (1995) and SBM (1995), the proposed Model I is a superior representation of consumer brand choice behavior for liquid detergents. Although the test fails to reject the model BV (1996) for the estimation sample ($p$-value $> 0.10$), it rejects it in favor of Model I for the holdout sample ($p < 0.01$). Thus, based on the Singleton (1985) test, we conclude that Model I performs better than the models of consideration set proposed by Andrews and Srinivasan (1995), Siddarth et al. (1995), and Bronnenberg and Vanhonacker (1996).

Table 3b reports two conventional measures of goodness of fit for the holdout sample of the liquid detergent data set: 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, AS (1995), and BV (1996) predict brand choice quite poorly. Although model SBM (1995) does a decent job, it does not predict choices as well as Model I.[24]

Traditional discrete choice models assume that displays and features influence brand choice by changing consumer preference (indirect utilities). On the other hand, in the proposed specification, we posit that the primary function of displays and features is to give costless price information to consumers about the posted price of the promoted brand. In our conceptualization, features and displays only affect brand choice probabilities by impacting the size and composition of the consideration set. To test whether displays/features affect the quality beliefs (or utilities) as assumed in the traditional brand choice models

[23] In the Singleton (1985) test, 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 $\chi^2$ with 1 d.f.

[24] Note that, unlike the competing models, Model I draws on demographics and other situational factors. Therefore, to show that Model I outperforms the other models not because it draws on more information than them, we have tested a variant of Model I that uses the same information as the competing models. We still find that the variant of Model I outperforms all the competing models for both liquid detergent and ketchup data sets.

**Table 3a** **Result from Comparison of Proposed Model (Model I) with Competing Models for Liquid Detergents [Model II, AS (1995), SBM (1995), and BV (1996)]**

| Model Comparison Using Singleton Test | Singleton Test Results | |
|---|---|---|
| | Estimation Sample | Holdout Sample |
| Model I against Model II | $J_S = 5.91(\chi^2$ with 1 d.f.) | $J_S = 6.84(\chi^2$ with 1 d.f.) |
| $H_0$: Model II is the "true" Model | $p$-value $< 0.02$ | $p$-value $< 0.01$ |
| $H_1$: Model I is the "true" Model | | |
| Model I against AS (1995) | $J_S = 3.67(\chi^2$ with 1 d.f.) | $J_S = 7.93(\chi^2$ with 1 d.f.) |
| $H_0$: AS (1995) is the "true" Model | $p$-value $< 0.10$ | $p$-value $< 0.01$ |
| $H_1$: Model I is the "true" Model | | |
| Model I against SBM (1995) | $J_S = 3.73(\chi^2$ with 1 d.f.) | $p$-value $< 0.10$ |
| $H_0$: SBM (1995) is the "true" Model | $J_S = 3.94(\chi^2$ with 1 d.f.) | $p$-value $< 0.05$ |
| $H_1$: Model I is the "true" Model | | |
| Model I against BV (1996) | $J_S = 2.53(\chi^2$ with 1 d.f.) | $p$-value $> 0.1$ |
| $H_0$: BV (1996) is the "true" Model | $J_S = 10.51(\chi^2$ with 1 d.f.) | $p$-value $< 0.01$ |
| $H_1$: Model I is the "true" Model | | |

**Table 3b** **Hit Rates and Likelihood Values for Models I, II, AS (1995), SBM (1995), and BV (1996) for Holdout Sample for Liquid Detergents**

| | Model I | Model II | AS (1995) | SBM (1995) | BV (1996) |
|---|---|---|---|---|---|
| Hit rate | 71.76% | 63.33% | 58.01% | 67.66% | 62.13% |
| Log-likelihood | 1,890 | 2,478 | 2,644 | 2,093 | 2,529 |

or affect the consumers' search behavior, we compare Model I to Model IA (Model IA is an alternative specification of the proposed model wherein features and displays only affect quality beliefs and not search costs).[25] Table TA.6 in the Technical Appendix presents the parameter estimates for Model IA. Table TA.7 in the Technical Appendix summarizes the Singleton results from comparison of Model I against Model IA for the liquid detergent data set for both the estimation and the holdout samples. We find that the test rejects Model IA in favor of Model I for both the estimation and holdout samples ($p$-value $< 0.05$). Table TA.8 in the Technical Appendix reports the hit rates and the log-likelihood values for both Models I and IA for the holdout samples. Again, we find that Model I outperforms Model IA both in terms of the log-likelihood and the hit rate. This supports our hypothesis that features and displays impact brand choice probabilities primarily through their influence on consumer search rather than through the utilities.

[25] We thank the area editor and an anonymous reviewer for this suggestion.

The parameter estimates for Model I for the ketchup data set are summarized in Table TA.9 of the Technical Appendix. For the ketchup data set, we compare the proposed model only to model SBM (1995; parameters are given in Table TA.10)—the best-performing alternative model for the liquid detergent data. Table TA.11 in the Technical Appendix summarizes the Singleton results from comparison of Model I with SBM (1995), whereas Table TA.12 reports the hit rates and the log-likelihood values for the holdout sample of ketchup for the two models. On all the three criteria, Model I outperforms SBM (1995) for both the estimation and the holdout samples.

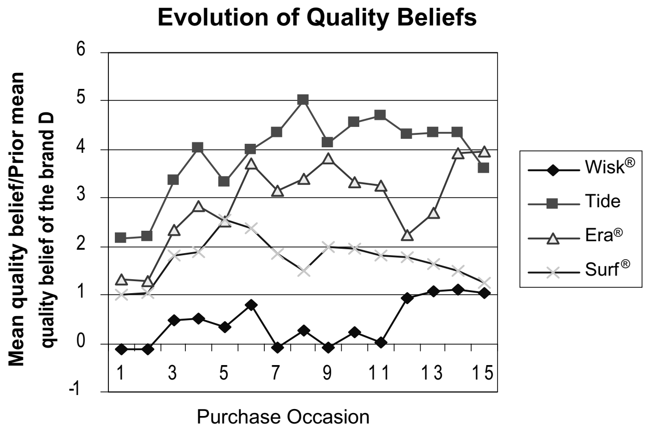
**Results for Liquid Detergents and Discussion of Managerial Insights.** We summarize our discussion under five broad substantive issues:

**Perceived Quality Positioning of the Competing Brands.** From Table 2, 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. The fact that Wisk has the lowest perceived quality is somewhat surprising because it happens to be the market share leader (31% share). This is because of the following two reasons: First, the average price of Wisk is the lowest among all the brands. Secondly, the average consideration set size when consumers buy Wisk is 2.21, which is significantly lower than the average consideration set size when consumers purchase other brands (Table 4). Also, when consumers buy Wisk, their consideration set is of size one 32% of the time. On the other hand, when consumers buy brands other than Wisk, their consideration set size is of size one 20–25% of the time (Table 4). It thus appears that many consumers buy Wisk out of inertia and that their purchase decision is not influenced by competitive promotions simply because they fail to notice the prices of the other brands. This is also consistent with the fact that Wisk is the most feature-advertised (frequency of feature ads = 0.14) and displayed (frequency of displays = 0.22) brand, and hence consumers incur much lower search costs for Wisk compared with other brands.[26]

**Consumer Learning and the Effectiveness of Quality Cues.** The estimate of $\alpha_0$—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 $\alpha_0$ suggests a low variance in subjective quality beliefs held by the consumers at the beginning of their consumption histories. Thus, a value of $\alpha_0$ implies that the extent of quality learning that can occur through consumption experience is low. Our estimate of $\alpha_0$ suggests a moderate extent of quality learning happening in the product category. Figure 2 shows the evolution of quality beliefs of the consumers across 15 purchase occasions. We can see that there is still a substantial amount of quality learning going on even after 15 consumption experiences. For Model II, the estimate of $\alpha_0$ 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 §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 et al. 1991). Thus, we conclude that ignoring consideration effects may lead to overestimation of consumer heterogeneity, thereby confirming the conclusions in Chiang et al. (1999).

[26] This is consistent with the finding of Fader and McAlister (1990), who show that consumers often restrict themselves to those brands that are heavily promoted.

**Figure 2** Evolution of Mean Quality Beliefs for Fifteen Purchase Occasions (Liquid Detergents)




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 $\kappa_f$ (the ratio of the noise-in-consumption signal to noise in the feature ad) and by $\kappa_d$ (the ratio of the noise-in-consumption signal to noise-in-display signal), respectively. High values of $\kappa_f$ and $\kappa_d$ imply that feature ads and displays have a strong impact on intrinsic preference. The parameter estimates of $\kappa_d = 0.041$ and $\kappa_f = 0.062$ suggest almost no impact of feature ads and displays on quality beliefs. This is expected because feature ads and displays primarily give price information (Blattberg and Neslin 1990).

**Quality-Price Trade-off and Consumer Price Sensitivity.** The estimates of the mean quality sensitivity parameter, $\overline{\theta}$, for Model I and Model 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 with low-price sensitivity). This makes sense

**Table 4 Consideration Set Sizes for Model I When the Respective Brands Are Chosen**

| Consideration Set Probabilities | Pr[No. of Brands = 1] | Pr[No. of Brands = 2] | Pr[No. of Brands = 3] | Pr[No. of Brands = 4] | Average Size (Std. Dev.) |
|---|---|---|---|---|---|
| Wisk chosen | 0.322 | 0.288 | 0.246 | 0.144 | 2.21 (0.031) |
| Tide chosen | 0.241 | 0.263 | 0.331 | 0.165 | 2.42 (0.036) |
| Era chosen | 0.252 | 0.278 | 0.421 | 0.049 | 2.27 (0.032) |
| Surf chosen | 0.208 | 0.329 | 0.367 | 0.096 | 2.35 (0.035) |

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 nonresponse 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 5a–5c). 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 (namely, Wisk and Tide). In contrast, the unconditional price elasticities for Model I are lower than the price elasticities for Model II for the low market share brands (namely, Era and Surf).[27]

**Table 5a Model I—Conditional Price Elasticities* 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 |

*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 5b Model I—Unconditional Price Elasticities for Liquid Detergents**

| Conditional 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 |

**Table 5c Model II—Price Elasticities for Liquid Detergents**

| Conditional 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 |

For the model AS (1995), 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-price vs. low-price tier) rather than the actual posted prices that determine consideration set composition. The estimates of the variance of quality sensitivity, $\sigma_\theta^2$, from Model I and Model II are 0.4251 and 1.219, respectively. This again

[27] We thank the area editor and an anonymous reviewer for alerting us to this distinction.

highlights the fact that ignoring heterogeneity in consideration sets may lead to overestimation of consumer heterogeneity as noted by Chiang et al. (1999).

**Determinants of the Costs of Price Search.** The estimate for the baseline search cost $C_0$ is 0.0472. Thus, the average size of the consideration set is 2.30 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. 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 6). This implies that a consumer is likely to buy a brand repeatedly out of inertia, and not from actively comparing the prices of all brands in this product category.

Table 6 also summarizes the marginal impact of feature ads and displays on the average 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 2.30 to 2.50 (the increase is statistically significant at $p < 0.01$). 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 baseline case of 2.30 to 2.75 (the increase is statistically significant at $p < 0.01$). 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 (namely, 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 $\alpha = 20\%$. The reduction in search cost during shopping trips on weekends compared with those during weekdays, $C_4$, is $-0.0032$ and is significant at $\alpha = 10\%$. Thus, the average size of consideration set during weekends is 2.34 and during the weekdays it is 2.30, thus suggesting a very modest impact of day of the week (*DOW*) on search behavior. This conforms to our expectation that consumers engage in more extensive price search

**Table 6** **Predicted Probabilities for Consideration Set Size with Varying Search Costs for Liquid Detergent Data Set**

| Specification of the Price Search Cost | Pr[No. of Brands = 1] | Pr[No. of Brands = 2] | Pr[No. of Brands = 3] | Pr[No. of Brands = 4] | Average Set Size (Std. Dev.) |
|---|---|---|---|---|---|
| $C_{ijt} = C_0^*$ | 0.234 | 0.320 | 0.353 | 0.091 | 2.30 (0.018) |
| $C_{ijt} = C_0 + C_1 \times DISPLAY_{ijt}$ (Display for all brands) | 0.212 | 0.273 | 0.315 | 0.198 | 2.50 (0.04) |
| $C_{ijt} = C_0 + C_2 \times FEATURE_{ijt}$ (Feature ad for all brands) | 0.182 | 0.214 | 0.266 | 0.336 | 2.75 (0.07) |
| $C_{ijt} = C_0 + C_4 \times DOW_{it}$ (Purchase is made on a weekend, $DOW = 1$) | 0.227 | 0.303 | 0.360 | 0.108 | 2.34 (0.023) |
| $C_{ijt} = C_0 + C_6 \times INCOME_i$ (Per capita household income of \$20,000) | 0.289 | 0.473 | 0.187 | 0.048 | 1.99 (0.08) |
| $C_{ijt} = C_0 + C_6 \times INCOME_i$ (Per capita household income of \$30,000) | 0.341 | 0.503 | 0.132 | 0.022 | 1.83 (0.10) |
| $C_{ijt} = 0$ (Price search is cost-less) | 0 (By assumption) | 0 (By assumption) | 0 (By assumption) | 1 (By assumption) | 4 |

*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.

during weekends because they have more time available for shopping (low opportunity cost of time).

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 2.30. The average consideration set size corresponding to per capita income of $20,000 and $30,000 are 1.99 and 1.83, respectively (the decrease in the average set size is statistically significant at $p < 0.01$). 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 $\alpha = 20\%$. Thus, previous displays do not significantly impact the costs of search during the current purchase occasion.

**Consumer Search Behavior.** Table 4 reports the frequency of the consideration set sizes for the proposed model when each of the four brands (namely, 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 despite its lower perceived quality, Wisk enjoys market share leadership primarily because a relatively high fraction of consumers who buy Wisk do not consider other brands.

Overall, we find that, in the liquid detergents category consumers consider one brand 26.2% of the time, two brands 28.7% of the time, three brands around 33.1% of the time, and four brands the rest (12)% of the time. This corresponds to an average consideration set size of 2.31. This indicates moderate intensity of price search in the product category. Table 7 compares the set-size probabilities predicted by the proposed model and the competing models: AS (1995), SBM (1995), and BV (1996). The average consideration set size for AS (1995), SBM (1995), and BV (1996) are 2.494, 2.564, and 2.585, respectively. We observe that the proposed model predicts a statistically significantly smaller consideration set. The reason we get lower consideration set size is as follows. The models proposed in Andrews and Srinivasan (1995), Siddarth et al. (1995), and Bronnenberg and Vanhonacker

**Table 7** **Consideration Set Sizes for Models I, AS (1995), SBM (1995), and BV (1996) for Liquid Detergents**

| Consideration Set Probabilities | Model I | AS (1995) | SBM (1995) | BV (1996) |
|---|---|---|---|---|
| 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 |
| Average Size | 2.31 | 2.494 | 2.564 | 2.585 |
| (Std. Dev.) | (0.03) | (0.033) | (0.026) | (0.029) |

(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.[28] Note that if we ignore intrinsic brand preference, we would get biased estimates of the distribution of consideration set sizes.

### 3.3. Cross-Category Analysis—Price Search Behavior in the Ketchup Category

The parameter estimates for the ketchup data set are given in Table TA.9 of the Technical Appendix. The parameters are by and large statistically significant and in the anticipated direction.

The estimate of mean quality sensitivity, $\bar{\theta}$, is 3.2140. This suggests that consumers are more price sensitive when purchasing liquid detergents than when purchasing ketchup. This is to be expected because detergents are a higher expense item than ketchup, and there is greater price variability in the detergent category compared with the ketchup category.[29] This may also suggest that through advertising, the leading brands in this category have been able to increase the

[28] For the proposed model, we recomputed the distribution of consideration set sizes setting the intercept term to zero. We find that, in this case, consumers consider one brand 19.7% of the time, two brands 27.9% of the time, three brands around 30.1% of the time, and four brands the rest (22.3%) of the time. Thus, the predicted average size of the consideration set is 2.55, which is comparable with those from AS (1995), SBM (1995), and BV (1996).

[29] We thank an anonymous reviewer for pointing this out.

intensity of preference for the differentiating attribute, thereby reducing price sensitivity. We find the variance of quality sensitivity, $\sigma_\theta^2 = 0.4739$, which is comparable with that for liquid detergents.

The value of the intrinsic search cost $C_0$ is estimated at 0.0284, which is lower than that for liquid detergents. This suggests that price search is less costly in the ketchup category. We conjecture that this could be because shelf space assigned to the ketchup category is much smaller than that assigned to liquid detergents, thereby making price comparisons easier for ketchup. Table 8a reports the average consideration set size of 2.25 under the baseline case where no brands are displayed or featured. Even though the baseline search cost for observing the price of a brand in the ketchup category is smaller than that in liquid detergents, we see a counterintuitive result that the average size of the consideration set is almost the same for both ketchup and detergents. The reason is that the variance of the prices of the brands is much higher on average in liquid detergents than in ketchup. This results in a larger propensity to search in liquid detergents than in ketchup, which further leads to larger consideration set sizes for liquid detergents. 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 the baseline case of 2.25 to 2.45. 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 $-0.0105$ and is significant. Thus, if all the brands are feature advertised, the average consideration set size increases from the baseline case of 2.25 to 2.38. The reduction in the search costs due to displays is higher in the ketchup category than in the liquid detergent category. Therefore, displays are more effective in giving price information to the consumers in the ketchup category than in the liquid detergent category.

**Table 8a Predicted Probabilities for Consideration Set Size with Varying Search Costs for the Ketchup Data Set**

| Specification for Price Search Costs | Size of the Consideration Set: Average Size (Std. Dev.) |
|---|---|
| $C_{ijt} = C_0^*$ | 2.25 (0.03) |
| $C_{ijt} = C_0 + C_1 \times DISPLAY_{ijt}$ (Display for all brands) | 2.38 (0.045) |
| $C_{ijt} = C_0 + C_2 \times FEATURE_{ijt}$ (Feature ad for all brands) | 2.45 (0.055) |
| $C_{ijt} = 0$ (Price search is costless) | 4 |

*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.

Similar to liquid detergents, in the ketchup category we find no significant impact of store familiarity and the presence of a full-time homemaker. Additionally, unlike liquid detergents, we find no significant impact of per capita income and the timing of purchase (weekend/weekday) on search costs.

Table 8b reports the frequency of consideration set sizes for the proposed model when each of the

**Table 8b Consideration Set Sizes for Model I When the Respective Brands Are Chosen for the Ketchup Data Set**

| Consideration Set Probabilities | Pr[# Brands = 1] | Pr[# Brands = 2] | Pr[# Brands = 3] | Pr[# Brands = 4] | Average Size (Std. Dev.) |
|---|---|---|---|---|---|
| Heinz chosen | 0.362 | 0.269 | 0.251 | 0.118 | 2.12 (0.037) |
| Hunt's chosen | 0.110 | 0.346 | 0.442 | 0.102 | 2.53 (0.033) |
| Del Monte chosen | 0.181 | 0.270 | 0.376 | 0.173 | 2.54 (0.033) |
| Generic chosen | 0.104 | 0.200 | 0.469 | 0.227 | 2.81 (90.031) |

four brands (namely, Heinz, Hunt's, Del Monte®, and generic) is chosen. We observe that the average size of the consideration set when Heinz was bought is 2.12, for Hunt's the average size is 2.53, for Del Monte® the average size is 2.54, and for the generic brand the average size is 2.81. As expected, consumers who buy the generic brand—being more price sensitive—undertake the most extensive search.

Overall, we find that, in the ketchup category, consumers consider one brand 28.7% of the time, two brands 27.6% of the time, three brands around 30.9% of the time and four brands the rest (12.9%) of the time. This corresponds to an average consideration set size of 2.28. This indicates that consideration sets for both liquid detergents and ketchup are almost the same.

## 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 because of frequent price promotions of varying depths of discount, there exists considerable price uncertainty; whereas 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 the posted prices 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.

Our model captures consumer heterogeneity in the following three distinct ways: first, the stochastic Bayesian quality learning captures the heterogeneity in the consumers' quality beliefs. Second, the random effects specification on the consumer's marginal willingness to pay for quality captures the slope heterogeneity, and third, letting consumers differ in their search costs on the basis of their demographic profile captures the heterogeneity in search costs. In our empirical analysis, we find that consumers incur significant search costs to discover the posted prices of the brands, which implies that they do not consider all the brands on a specific shopping trip. Further, we find that, for the two product categories of interest, although in-store display activities and feature ads do not influence consumers' quality perceptions of the brands, they significantly reduce the search costs of the brands and thereby increase the probability of the brands being considered. Also, we find that in-store feature ads reduce the consumers' search costs more than in-store displays in both the product categories, and per capita income of the consumer's household significantly increases their search costs in the liquid detergent product category. Finally, while comparing our model with other competing models on consideration, we show that the consumers' price sensitivity is seriously underestimated if we were to assume that consumers get to know the posted prices of all the brands at zero cost, and the consideration set sizes can be seriously overestimated (as in other competing models of consideration) if we do not include the brand intercept term in the utilities in the consideration stage.

We consider this paper an important first step to in studying 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. First, in our specification, as the number of alternatives increase, the combinatorial space of the number of consideration sets explodes, and the estimations become much more challenging. Nevertheless, having a large number of alternatives does not make it impossible to implement the model because we use the MSM methodology that considerably eases the estimation. Second, 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. Third, we assume that the consumers know the true price distributions of all the brands. This assumption can be relaxed if we were to assume that consumers learn about the price distributions of brands from the observed posted prices. Fourth, although we recognize consumer uncertainty about brand qualities and model evolution of consumer quality perceptions, because 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 multiarmed bandit problem). Fifth, we have used "brand composites" rather than stock keeping units (SKUs) as the discrete alternatives that consumers consider and select. In other words, all the UPC items sharing the same brand name are aggregated to create one single brand. However, if the analysis were done at the UPC level, several UPC items would share the same brand name. Therefore, even if a UPC item is never purchased, the consumer would still learn about its quality because of correlated quality learning. In that case, we would have to assume a correlated quality learning mechanism (Erdem 1998). Finally, our model raises the possibility that a brand with different sizes can influence the inclusion of one or more sizes of the brand in the consumer's consideration set by changing the prices of its offerings or selective promotion through feature or display. Such a strategic behavior may be an added reason for brand size/flavor proliferation. An analysis at the brand-size level will offer additional insights in to the issue.[30] Finally, Gonul and Srinivasan (1996) show that future promotion expectations might affect current purchase behavior. Incorporating the implications for current consideration set is another significant issue. We hope to address these issues in a future research effort.

## Acknowledgments
The authors thank the editor, the area editor, and three anonymous reviewers for their constructive comments and suggestions. They also benefited from insightful comments from the seminar participants at the University of Toronto, the University of Florida at Gainesville, Columbia University, New York University, Stanford University, and the University of Chicago. This work was partly supported by the Cannaught Grant (at Rothman School of Management, University of Toronto) to the first author and Beatrice Companies Faculty Research Fund (at the Graduate School of Business, University of Chicago) to the second author. Please forward any correspondence to the third author. The usual disclaimer applies.

[30] We thank an anonymous reviewer for pointing this out.

## Appendix
Proposition 1. *The greater the value of the variance of the utility function in the consideration stage,* $\sigma^2_{u_{ij,t}} = \sigma^2_{p_j}$, *the greater is the size of the optimal consideration set.*

Proof. Consider the case in which the variance of the prices of all the brands is the same (i.e., $\sigma^2_{p_j} = \sigma^2_p$ for all brands $j$). Define $\beta = \sqrt{6}\sigma_p/\pi$ and $\overline{u}_{ij,t} = \theta\omega_{ij,t} - \overline{P}_j$. Also, define the expected benefit for choosing any consideration set $\{h\}$ as $EB_{\{h\}} = E\max[\{u_{ik,t}\}_{k\in\{h\}} - \sum_{k\in\{h\}} C_{ik,t}]$. 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 $\beta$ 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 in the expected benefits for considering the consideration sets $\{j\}$ and $\{\tilde{j}\}$

$$\begin{aligned} EB_{\{\tilde{j}\}} - EB_{\{j\}} &= \beta \log\Big(\exp(\overline{u}_{il,t}/\beta) + \sum_{k\in\{j\}} \exp(\overline{u}_{ik,t}/\beta)\Big) \\ &\quad - C_{il,t} - \beta\log\Big(\sum_{k\in\{j\}} \exp(\overline{u}_{ik,t}/\beta)\Big) \\ &= \beta \ln\left(1 + \frac{\exp(\overline{u}_{il,t}/\beta)}{\sum_{k\in\{j\}} \exp(\overline{u}_{ik,t}/\beta)}\right) - C_{il,t}. \end{aligned} \tag{A1}$$

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

## References
Allenby, G. M., J. L. Ginter. 1995. The effects of in-store displays and feature advertising on consideration sets. *Internat. J. Res. Marketing* **12**(1) 67–81.
Andrews, R. L., T. C. Srinivasan. 1995. Studying consideration effects in empirical choice models using scanner panel data. *J. Marketing Res.* **32**(1), 30–41.
Bagwell, K., M. Riordan. 1991. High and declining prices signal product quality. *Amer. Econom. Rev.* **81** 224–239.
Blattberg, R. C., S. A. Neslin. 1990. *Sales Promotions: Concepts, Methods and Strategies*. Prentice-Hall, New York.

——, C. J. Wisniewski. 1989. Price induced patterns of competition. *Marketing Sci.* **8**(4) 291–299

Bronnenberg, B. J., W. R. Vanhonacker. 1996. Limited choice sets, local price response, and implied measures of price competition. *J. Marketing Res.* **33**(2) 163–173.

Bucklin, R. E., J. M. Lattin. 1991. A two-state model of purchase incidence and brand choice. *Marketing Sci.* **10**(1) 24–39.

Chiang, J., S. Chib, C. Narasimhan. 1999. Markov chain Monte Carlo and models of consideration set and parameter heterogeneity. *J. Econometrics* **89** (Mar/Apr) 223–248.

Chintagunta, P. K. 1993. Estimating a multinomial probit model of brand choice using the method of simulated moments. *Marketing Sci.* **11**(4) 386–407.

——, D. Jain, N. Vilcassim. 1991. Investigating heterogeneity in brand preference in logit model for panel data. *J. Marketing Res.* **28** 417–428

Cyert, Richard M., Morris DeGroot. 1974. Rational expectations and Bayesian analysis. *J. Political Econom.* **82**(3) 521–536.

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

Desai, K. K., W. D. Hoyer. 2000. Descriptive characteristics of memory-based consideration sets: Influence of usage occasion frequency and usage location familiarity. *J. Consumer Res.* **27**(3) 309–323.

Easley, David, Nicholas M. Keifer. 1988. Controlling a stochastic process with unknown parameters. *Econometrica* **56**(5) 1045–1064.

Erdem, T. 1998. An empirical analysis of umbrella branding. *J. Marketing Res.* **35**(3) 339–352.

——, M. P. Keane. 1996. Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Sci.* **15**(1) 1–20.

Fader, P. S., L. McAlister. 1990. An elimination by aspects model of consumer response to promotions. *J. Marketing Res.* **27**(3) 322–332.

Feinberg, F. M., J. Huber. 1996. A theory of cutoff formation under imperfect information. *Management Sci.* **42**(1) 65–84.

Gensch, Dennis H. 1987. A two stage disaggregate attribute choice model. *Marketing Sci.* **6**(2) 223–240.

Gonul, F., K. Srinivasan. 1993. Modeling multiple sources of heterogeneity in multinomial logit models: Methodological and managerial issues. *Marketing Sci.* **12**(3) 213–229.

——, ——. 1996. Estimating the impact of consumer expectations of coupons on purchase behavior: A dynamic structural model. *Marketing Sci.* **15**(3) 262–279.

Guadagni P. M., J. D. C. Little. 1983. A logit model of brand choice calibrated on scanner data. *Marketing Sci.* **2**(3) 203–238.

Hanemann, W. M. 1984. Discrete/continuous models of consumer demand. *Econometrica* **52**(3) 541–562.

Hauser, J. R., B. Wernerfelt. 1990. An evaluation cost model of consideration sets. *J. Consumer Res.* **16**(March) 393–408.

Horstmann, I., G. M. MacDonald. 1994. When is advertising a signal of product quality? *J. Econom. Management Strategy* **3**(Fall) 561–584.

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

Johnson, N., S. Kotz. 1974. *Distributions in Statistics—Continuous Multivariate Distributions*. John Wiley and Sons, New York.

Jovanovic, Boyan. 1979. Job matching and theory of turnover. *J. Political Econom.* **87**(5) 972–990.

Kalra, A., S. Rajiv, K. Srinivasan. 1998. Response to competitive entry: A rationale for delayed defensive reaction. *Marketing Sci.* **17**(4) 380–405.

Kamakura, W. A., G. J. Russell. 1989. A probabilistic choice model of market segmentation and elasticity structure. *J. Marketing Res.* **26**(4) 379–391.

Kotler, P. 1996. *Marketing Management: Analysis, Planning, Implementation and Control*, 9th ed. Prentice-Hall, New York.

Manrai, A. K., R. L. Andrews. 1998. Two-stage discrete choice models for scanner panel data: An assessment of process and assumptions. *Eur. J. Oper. Res.* **111**(20) 193–215.

McFadden, D. 1981. Econometric models of probabilistic choice. C. Manski, D. McFadden, eds. *Structural Analysis of Discrete Data with Econometric Applications*. MIT Press, Cambridge MA.

——. 1989. A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica* **57** 995–1026.

Mehta, Nitin. 2000. Essays in consumer search behavior. Unpublished dissertation, Graduate School of Industrial Administration, Carnegie Mellon University, Philadelphia, PA.

Milgrom, P., J. Roberts. 1986. Prices and advertising signals of product quality. *J. Political Econom.* **94** 796–821.

Miller, Robert A. 1984. Job matching and occupation choice. *J. Political Econom.* **92**(6) 1086–1120.

Mitra, A. 1995. Advertising and the stability of consideration sets over multiple purchase occasions. *Internat. J. Res. Marketing* **12**(1) 81–94.

Moorthy, K. S. 1988. Product and price competition in a duopoly. *Marketing Sci.* **7**(2) 141–168.

——, J. Banks. 1999. A model of price promotions with consumer search. *Internat. J. Indust. Organ.* **17** 371–398.

——, K. Srinivasan. 1995. Signaling quality with a money-back guarantee: The role of transaction costs. *Marketing Sci.* **14**(1) 442–466.

Murthi, B. P. S., K. Srinivasan. 1999. Consumer's extent of evaluation on brand choice. *J. Bus.* **72**(2) 229–256.

Narasimhan, C. 1988. Competitive promotional strategies. *J. Bus.* **61**(October) 427–449.

Nedungadi, P. 1990. Recall and consumer consideration sets: Influencing choice without altering evaluation. *J. Consumer Res.* **17**(December) 263–276.

Nelson, P. 1974. Advertising as information. *J. Political Econom.* **81** 729–754.

Newman, J. W., R. Staelin. 1972. Pre-purchase information seeking for new cars and major household appliances. *J. Marketing Res.* **9**(August) 249–257.

Pakes, A., D. Pollard. 1989. Simulation and the asymptotics of optimization estimators. *Econometrica* **57**(5) 1027.

Raju, J. S., V. Srinivasan, R. Lal. 1990. The effects of brand loyalty on competitive price promotional strategies. *Management Sci.* **36**(March) 276–304.

Ratchford, B. 1982. Cost-benefit models for explaining consumer choice and information seeking behavior. *Management Sci.* **28**(2) 197–212.

Robert, J., D. O. Stahl. 1993. Informative price advertising in a sequential search model. *Econometrica* **61**(3) 657–686.

Roberts, J. H., J. M. Lattin. 1991. Development and testing of a model of consideration set composition. *J. Marketing Res.* **28**(4) 429–440.

——, J. Urban. 1988. Modeling multi-attribute utility risk and belief dynamics. *Management Sci.* **34**(2) 167–185.

Roy, R., P. K. Chintagunta, S. Haldar. 1996. A framework for analyzing habits, hand-of-past, and heterogeneity in dynamic brand choice. *Marketing Sci.* **15**(3) 280–299.

Seetharaman, P. B., A. Ainslie, P. K. Chintagunta. 1999. Investigating household state dependence effects across categories. *J. Marketing Res.* **36**(4) 488–500.

Siddarth, S., R. E. Bucklin, D. G. Morrison. 1995. Making the cut: Modeling and analyzing choice set restrictions in scanner panel data. *J. Marketing Res.* **32**(8) 255–266.

Singleton, K. J. 1985. Testing specifications of economic agents intertemporal optimum problems in the presence of alternative models. *J. Econometrics* **30** 391–413.

Stigler, G. J. 1961. The economics of information. *J. Political Econom.* **69** 213–235.

Villas-Boas, J. M. 1998. Product line design for a distribution channel. *Marketing Sci.* **17**(2) 156–169.

Wernerfelt, B. 1991. Brand loyalty and market equilibrium. *Marketing Sci.* **10**(3) 229–245.

Willig, R. D. 1978. Incremental consumers' surplus and hedonic price adjustment. *J. Econom. Theory* **17** 227–253.

*This paper was received June 12, 2001, and was with the authors 4 months for 2 revisions; processed by Greg Allenby.*