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

Nitin Mehta
Surendra Rajiv
Kannan Srinivasan

PRICE UNCERTAINTY AND CONSUMER SEARCH:
A STRUCTURAL MODEL OF CONSIDERATION SET FORMATION

TECHNICAL SUPPLEMENT

CONTENTS

1. NOTATIONS

SYMBOL	EXPLANATION
$U_{ij,t}$	Indirect utility or consumer surplus associated with brand *j* at time *t* for household *i*.
q_j	"True" quality of brand *j*. This is assumed to be the same for all consumers
$q_{ij,t}$	Perceived quality of brand *j* at time *t* for household *i*. The perceived quality is a normally distributed random variable which converges to the "true" quality as the consumer learns over her purchase history i.e. $q_{ij,t} \to q_j$ as $t \to \infty$
$p_{ij,t}$	Price of brand *j* at time *t* for the household *i*. Since prices fluctuate over time due to sales promotions, this is a random variable till consumer discovers the "actual" posted price. Said differently, $p_{ij,t}$ is a random variable for the consumer at the consideration stage, but it is a deterministic variable (being the posted price of brand *j* at time *t*) at the choice stage.
q	The quality sensitivity of a randomly drawn household in the population. This is assumed to be a random draw from a gamma distribution with mean $\bar{q}$ and variance s_q^2. This is also assumed to be the same across all brands for a given household.
$\bar{q}$	The mean quality sensitivity across all households in the population. This is assumed to be the same across all brands.
s_q^2	The variance in the quality sensitivity across the entire household population.
$w_{ij,t}$	The mean of the perceived quality belief of brand *j* at time *t* for the household *i*.
$\tilde{w}_{ij,t}$	The analyst's probabilistic estimate of the mean of the perceived quality belief of brand *j* at time *t* for the household *i*.
$s_{w_{ij,t}}^2$	The variance of the perceived quality belief of brand *j* at time *t* for the household *i*.
$l_{ij,t-1}$	Quality signal about brand *j*'s quality received by household *i* from consuming the brand *j* at time *t*-1. This signal is assumed to be "noisy" so that $l_{ij,t-1} \sim N\left(q_j, s_h^2\right)$. Note that the consumer observes the realized value of this signal; however, from the analyst's perspective, it is an unobservable.
$h_{ij,t-1}$	"Noise" in the quality signal received about brand *j* after its consumption at time *t*-1 by the household *i*. Thus, $l_{ij,t-1} = q_j + h_{ij,t-1}$.
s_h^2	Variance of the consumption signal that is assumed to be the same across all households and across all brands and across all purchase occasions.
$a_{ij,t}$	$a_{ij,t} = s_h^2 \big/ s_{w_{ij,t}}^2$ is ratio of the variance of the (noisy) consumption signal to the variance of the perceived quality of brand *j* at *t* for the household *i*. Since we normalize $s_h^2 = 1$ for parameter identification, $a_{ij,t}$ can be interpreted as the inverse of the uncertainty in the quality of brand *j*, at time *t*, for the household *i*.
a_0	Value of the noise to information ratio, $a_{ij,t}$, at t = 0. This is assumed to be the same for all brands and for all households.

SYMBOL	EXPLANATION
$\bar{p}_j$	Mean of the price distribution of brand j. This is assumed to be the same across all time periods and across all households.
$s^2_{p_j}$	Variance of the price distribution of brand j. This is assumed to be the same across all time periods and across all households.
$s^2_{u_{ij,t}}$	Variance of the indirect utility of brand j at the consideration stage on household i's tth purchase occasion.
e_c	Euler's constant.
$C_{ij,t}$	Search cost for ascertaining price of brand j on purchase occasion t for household i.
$d_{ij,t-1}$	Indicator variable such that $d_{ij,t-1}$ = 1 if household i purchases brand j on purchase occasion t-1.

2. MODEL FORMULATION – ANALYTICAL DETAILS

2.1 Evolution of the Quality Perception for the Proposed Model:

We assume that the perceived quality of brand *j* on purchase occasion *t* for the household i, $q_{ij,t}$, is a normal random variable with mean $w_{ij,t}$ and variance $s^2_{w_{ij,t}}$. Note that the mean $w_{ij,t}$ need not coincide with the "true" quality of brand j, viz., q_j (although over their purchase history, $w_{ij,t}$ will converge to q_j). This captures the fact that consumers are uncertain about the qualities of the different brands in the product category and only hold beliefs about their qualities. Thus, under these assumptions

$$q_{ij,t} \sim N\left(w_{ij,t}, s^2_{w_{ij,t}}\right). \qquad \text{(A2.1)}$$

Further, we assume consumer can learn about the "true" qualities of brand j, q_j, through the consumption experiences on prior purchase occasions. However, we assume that this mechanism only provides "noisy" signals of the "true" quality. The direct implication of this assumption is that the "true" quality does not get revealed completely after just one consumption-experience. We operationalize consumer learning about brand qualities through "noisy" consumption signals as follows.

Let $l_{ij,t-1}$ denote the quality cue associated with consumption experience in time *t-1* by consumer *i* about brand *j* specified as follows:

$$l_{ij,t-1} = q_j + h_{ij,t-1}. \qquad \text{(A2.2)}$$

In equation (A2.2), q_j denotes the "true" quality of brand j. Further, $h_{ij,t-1}$ denotes the noise associated with the consumption signal, $l_{ij,t-1}$. For simplicity, we assume that $h_{ij,t-1}$ is i.i.d across all consumers, across all brands and across all purchase occasions.

To exploit the self-conjugacy of the normal density, we assume that $h_{ij,t-1}$ is a normal random variable with mean equal to zero and variance equal to s^2_h, i.e.,

$$h_{ij,t-1} \sim N\left(0, s^2_h\right) \qquad \text{(A2.3)}$$

Thus, s^2_h is a measure of the non-informativeness of the consumption experience and if either s^2_h = 0, the consumer will get to learn the "true" brand quality from just one consumption experience.

Consider consumer *i* who might receive consumption signal about the quality of brand *j* between *t*-1 and *t* purchase occasions. Specifically, let $d_{ij,t-1}$ be the indicator variable such that

$d_{ij,t-1}$	=	1	if consumer *i* purchases brand *j* on purchase occasion at *t*-1 (i.e., the consumer receives the consumption signal before the purchase occasion *t*); and,
	=	0	otherwise.

Consider the case where $d_{ij,t-1} = 1$ and the consumer i receives the realization of the random consumption signal $\hat{l}_{ij,t-1}$ at time *t*-1. Before receiving the quality signal $\hat{l}_{ij,t-1}$, the consumer would have prior subjective beliefs about the quality of brand j, based on her purchase history at time t-1, $\mathrm{H}_i(t-1)$. Let the prior beliefs of the consumer *i* be denoted by

$$q_{ij,t} \mid H_i(t-1) \quad \sim \quad N\left(w_{ij,t-1}, s^2_{w_{ij,t-1}}\right). \qquad \text{(A2.4)}$$

After observing the realization of the random consumption signal $\hat{l}_{ij,t-1}$, the consumer would update her subjective beliefs about the quality of brand j in a Bayesian fashion. Represent her posterior beliefs, after observing the quality cue $\hat{l}_{ij,t-1}$, by

$$q_{ij,t} \mid H_i(t) \;\sim\; N\left(w_{ij,t}, s^2_{w\,ijt}\right) \tag{A2.5}$$

Then, because of the self-conjugacy of the normal density, the mean and variance of the posterior beliefs are related to the mean and variance of the prior beliefs as follows (DeGroot, 1970):

$$w_{ij,t} = \frac{\dfrac{w_{ij,t-1}}{s^2_{w_{ij,t-1}}} + d_{ij,t-1}\dfrac{\hat{l}_{ij,t-1}}{s^2_h}}{\dfrac{1}{s^2_{w_{ij,t-1}}} + d_{ij,t-1}\dfrac{1}{s^2_h}}, \text{ and,} \tag{A2.6}$$

$$\frac{1}{s^2_{w_{ij,t}}} = \frac{1}{s^2_{w_{ij,t-1}}} + d_{ij,t-1}\frac{1}{s^2_h}. \tag{A2.7}$$

By defining $a_{ij,t-1} = s^2_h \big/ s^2_{w_{ij,t-1}}$, we can rewrite equations (A2.6) and (A2.7) as follows:

$$w_{ij,t} = \frac{w_{ij,t-1}a_{ij,t-1} + d_{ij,t-1}\hat{l}_{ij,t-1}}{a_{ij,t-1} + d_{ij,t-1}}, \text{ and,} \tag{A2.8}$$

$$a_{ij,t} = a_{ij,t-1} + d_{ij,t-1}. \tag{A2.9}$$

Equation (A2.8) 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. Note that the mean of the consumer's quality beliefs converges to the true quality q_j as $t \to \infty$. In other words, after the consumer has observed the infinite sequence of random quality draws $\left\{\hat{l}_{ij,t}\right\}_{t=1}^{t=\infty}$, her mean quality beliefs of brand j will converge to the true quality q_j. This can be shown by recursively substituting $l_{ij,s} = q_j + h_{ij,s}$ for all s=0 till t-1 into equation (A2.8) to get the following expression:

$$w_{ij,t} = q_j\left(\frac{\sum_{s=1}^{t-1} d_{ij,s}}{a_0 + \sum_{s=1}^{t-1} d_{ij,s}}\right) + q_j\left(\frac{\sum_{s=1}^{t-1} d_{ij,s}h_{ij,s}}{a_0 + \sum_{s=1}^{t-1} d_{ij,s}}\right) \tag{A2.10}$$

Using Law of Large Numbers, we can show that $\lim_{t\to\infty}\sum_{s=1}^{t} d_{ij,s}h_{ij,s} \to 0$ and hence $w_{ij,t} \to q_j$ as $t \to \infty$.

It is important to note that the consumer observes the realization of the random signals $\left\{\hat{l}_{ij,s}\right\}_{s=1}^{t-1}$ and hence, she knows deterministically the values of $w_{ij,t}$ for all brands j at any time t. But the analyst does not observe the realizations of the random quality draws $\left\{\hat{l}_{ij,s}\right\}_{s=1}^{t-1}$. From the analyst's perspective, $\left\{\hat{l}_{ij,s}\right\}_{s=1}^{t-1}$ are random variables and hence, the analyst's estimate of $w_{ij,t}$ (which we denote by $\tilde{w}_{ij,t}$) is a random variable too. Since the analyst's information set at time t consists of the sequence of the choices made by the consumer from time s=1 till time s=t-1, he can use this information to get a more precise estimate of the mean quality beliefs of all brands at time t. This is because the sequence of past choices of consumer i $\left\{d_{i1,s}, d_{i2,s}..d_{iN,s}\right\}_{s=1}^{t-1}$ imposes a set of restrictions on the state space of her past utilities and hence on her past quality beliefs $\{\tilde{w}_{i1,s}, \tilde{w}_{i2,s} .. \tilde{w}_{iN,s}\}_{s=1}^{t-1}$. For instance, if brand k was purchased by consumer i in time period s (where $s<t$), a set of restrictions (represented as $\{R_{i,s}\}$ for future reference) will be imposed on the state space of the mean quality beliefs $\{\tilde{w}_{i1,s}, \tilde{w}_{i2,s} .. \tilde{w}_{iN,s}\}$ at time s. Since Bayesian learning implies that the mean quality beliefs are correlated across time, it follows that the mean quality beliefs of all brands at time t will also have restrictions $\{R_{i,s}\}_{s=1}^{t-1}$ imposed on their state space. These restrictions are discussed in greater detail in Section 3.1.

From the analyst's perspective at time t, $l_{ij,t-1}$ will be a random variable as given in equation (A2.2) by $l_{ij,t-1} \sim N\left(q_j, s_h^2\right)$. But note that at time t, neither the consumer nor the analyst knows the true quality q_j of ant brand j. Hence, the analyst can not assume q_j to be the mean of the quality signal $l_{ij,t-1}$. The analyst can only use his most recently updated belief (as per his information set at time t) about the true quality q_j to be the mean of the signal $l_{ij,t-1}$. In order to know the analyst's belief of q_j at time t, it will be instructive to look at the elements in the analyst's and the consumer's information set at time t. The elements in the information sets of the consumer and the analyst are explained in Figure 1.

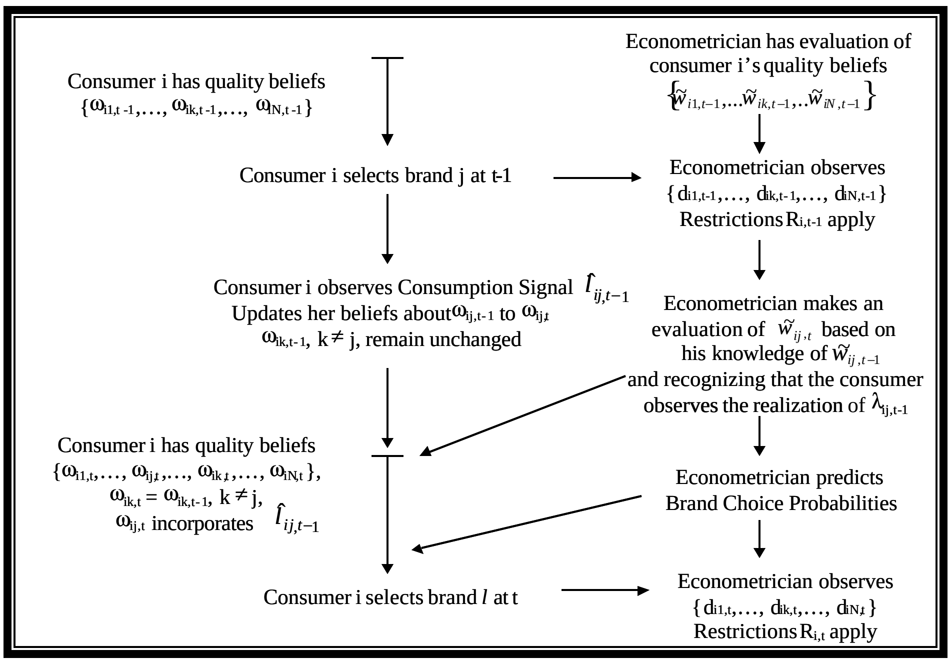

FIGURE 1: Sequence of Events and Information Sets

Note that in Figure 1 while the consumer observes the realization $\hat{l}_{ij,t-1}$, the analyst does not. Further note that the consumer gets to observe the realization $\hat{l}_{ij,t-1}$ *after* the analyst observes her brand choice in period t-1 $\{d_{i1,t-1}, \ldots, d_{ik,t-1}, \ldots, d_{iN,t-1}\}$. As we pointed out before, the observed brand choice at t-1 helps the analyst to make a more precise estimate of the mean of consumer's quality beliefs at t-1 i.e., $\{\tilde{w}_{i1,t-1}, \tilde{w}_{i2,t-1} .. \tilde{w}_{iN,t-1}\}$ by imposing the restrictions $R_{i,t-1}$. However, the analyst does not get any additional information about $\hat{l}_{ij,t-1}$ till the consumer makes brand choice at t (i.e., when restrictions $R_{i,t}$ kick in). Thus, for him, $l_{ij,t-1}$ is a normal random variable with "no restrictions on its state space." This also implies that as far as $l_{ij,t-1}$ is concerned, the analyst's information set at the time of predicting brand choice at t **is identical** to consumer's information set at t-1 before she observes $\hat{l}_{ij,t-1}$. Said differently, if the consumer were to evaluate her mean quality belief at time t, before having observed $\hat{l}_{ij,t-1}$, her treatment of $l_{ij,t-1}$ and hence her estimate of $\tilde{w}_{ij,t}$ would be no different than the analyst's evaluation. Therefore, conditional on the value of the mean quality belief at time t-1 (viz., $\tilde{w}_{ij,t-1}$), the analyst assumes the distribution as $l_{ij,t-1} \sim N\left(q_{ij,t-1}, s_h^2\right)$ where $q_{ij,t-1}$ is the most recent belief of the consumer about the quality of brand j just before receiving the signal $\hat{l}_{ij,t-1}$. Since $q_{ij,t-1}$ itself is normally distributed with $q_{ij,t-1} \sim N\left(\tilde{w}_{ij,t-1}, a_{ij,t-1}^{-1}\right)$, the signal $l_{ij,t-1}$ will be distributed as $l_{ij,t-1} \sim N\left(\tilde{w}_{ij,t-1}, s_h^2 + a_{ij,t-1}^{-1}\right)$. Further since we set $s_h^2 = 1$ for identification purposes, the signal $l_{ij,t-1}$ received at time t-1 will be distributed as $l_{ij,t-1} \sim N\left(\tilde{w}_{ij,t-1}, 1 + a_{ij,t-1}^{-1}\right)$ from the analyst's perspective at time t. Substituting $l_{ij,t-1} \sim N\left(\tilde{w}_{ij,t-1}, 1 + a_{ij,t-1}^{-1}\right)$ in equation (A2.9) we get the analyst's estimate of $\tilde{w}_{ij,t}$ conditional on $\tilde{w}_{ij,t-1}$ as

$$\tilde{w}_{ij,t} = \tilde{w}_{ij,t-1} + N\left(0, \frac{d_{ij,t-1}\left(1 + a_{ij,t-1}^{-1}\right)}{\left(a_{ij,t-1} + d_{ij,t-1}\right)^2}\right) \qquad \text{(A2.11)}$$

Note that our approach outlined above for deriving the analyst's beliefs of $\tilde{w}_{ij,t}$ (conditional on $\tilde{w}_{ij,t-1}$) is consistent with Jovanovic (1979) and Miller (1984).[5] We can see from equation (A2.11) that conditional on $\tilde{w}_{ij,t-1}$, the analyst's *ex ante* evaluation of the mean of consumer's quality belief at time $r>t$ implies that $E\left[\tilde{w}_{ij,r} \mid \tilde{w}_{ij,t-1}\right] = \tilde{w}_{ij,t-1} \forall r > t-1$. This implies that the analyst's ex ante estimate of the mean of the quality beliefs of brand j (conditional on $\tilde{w}_{ij,t-1}$) converges to $\tilde{w}_{ij,t-1}$ as $t \to \infty$. This is not to say that *ex post* the analyst's evaluation of the mean of consumer's quality belief does not converge to q_j as $t \to \infty$. It does because the analyst observes the sequence of consumer choices $\{d_{i,s}\}_{s=1}^{s=t-1}$ that imposes restrictions $\{R_{i,s}\}_{s=1}^{t-1}$ on $\tilde{w}_{ij,t-1}$ as discussed before. As $t \to \infty$, the number of such restrictions on $\tilde{w}_{ij,t-1}$ becomes infinite and hence the analyst's *ex post* evaluation (after observing $\{d_{i,s}\}_{s=1}^{s=t-1}$) converges to the point q_j.

Also note that equation (A2.11) gives the analyst's estimate of $\tilde{w}_{ij,t}$ conditional on $\tilde{w}_{ij,t-1}$. In order to get the unconditional estimate of $\tilde{w}_{ij,t}$ from the analyst's perspective, we can recursively derive the expression of $\tilde{w}_{ij,t-1}$ as a function of the prior mean quality belief $\tilde{w}_{ij,0}$ of brand j at time 0 and the truncated normal random variables ($\text{Trunacted}\left\{N\left(0, \frac{d_{ij,s}\left(1+a_{ij,s}^{-1}\right)}{\left(a_{ij,s-1}+d_{ij,s}\right)^2}\right)\right\}_{s=1}^{t-2}$) that get added due to consumption of brand j from time s=1 till time s=t-2. Note that the random variables $\left\{N\left(0, \frac{d_{ij,s}\left(1+a_{ij,s}^{-1}\right)}{\left(a_{ij,s-1}+d_{ij,s}\right)^2}\right)\right\}_{s=1}^{t-2}$ will be

[5] Jovanovic, Boyan (1979): "Job Matching and Theory of Turnover," *Journal of Political Economy*, Vol. 87, No. 5, 972-990; Miller, Robert A. (1984): "Job Matching and Occupation Choice," *Journal of Political Economy*, Vol. 92, No. 6, 1086-1120.

truncated as a result of the restrictions $\{R_{i,s}\}_{s=1}^{t-1}$ imposed on them because of the choices made by the consumer i till time t-1.

For estimation purposes, we assume that the means of the quality beliefs for all brands and for all the consumers at the beginning of their consumption history are zero. Also, we assume that the value of the precision of the subjective quality beliefs of all the brands for all the consumers at the beginning of their consumption history is a_0. In other words, $\tilde{w}_{ij,0} = 0$ and $a_{ij,0} = a_0$ for all brands j and for all consumers i. We can then recursively derive the expressions for $\tilde{w}_{ij,t}$ and $a_{ij,t}$ as a sum of all previous consumption signals as follows:

$$\tilde{w}_{ij,t} = N\left(0, \frac{d_{ij,t-1}\left(1+a_{ij,t-1}^{-1}\right)}{\left(a_{ij,t-1}+d_{ij,t-1}\right)^2}\right) + \sum_{s=1}^{t-2} \underset{\{R_{i,s}\}_{s=1}^{t-1}}{\text{Truncated}}\left[N\left(0, \frac{d_{ij,s}\left(1+a_{ij,s}^{-1}\right)}{\left(a_{ij,s}+d_{ij,s}\right)^2}\right)\right], \text{ and,} \qquad \text{(A2.12)}$$

$$a_{ij,t} = a_0 + \sum_{s=1}^{t} d_{ij,s-1} \qquad \text{(A2.13)}$$

The first expression on the right hand side of equation (A2.12) is a normally distributed random variable. As discussed before, it is not subject to any restrictions as a result of the past sequence of choices. It captures the effect of the consumption signal received between time t-1 and t on the mean of the perceived quality of brand j. Note that the realizations of $l_{ij,t-1}$ can take both positive as well as negative values. Thus, unlike the reduced-form operationalization of state dependence/brand loyalty (e.g., Guadagni and Little, 1983), the proposed model allows for both upward and downward shift in the consumer's intrinsic brand preference over her purchase history.

2.2 Formation of the Optimal Consideration Set for the Proposed Model -- Consumer's Problem:

We assume that consumer i's (indirect) utility from brand j on purchase occasion t can be approximated as a linear function of brand j's perceived quality, $q_{ij,t}$, and price, p_{ijt}, as follows:

$$U_{ijt} = q q_{ijt} - p_{ijt}, \qquad \text{(A2.14)}$$

The parameter q denotes the consumer's intrinsic preference for quality. This is assumed to be gamma distributed across the households with mean $\bar{q}$ and variance s_q^2. For simplicity, we assume q to be the same across the brands for a given consumer.

Based on the discussion in Section 2.1, the quality of the various brands, $q_{ij,t}\ \forall j$, is a random variable. Further, at the consideration stage, consumer i does not know the posted price of any brand. The consumer first decides the brands whose posted prices on that purchase occasion she will search. The set of brands whose prices she searches is characterized as her consideration set.

Note that prior to engaging in price search, the consumer is only aware of the distributions of the prices of all the brands in the product category; as such, the actual posted price of the brands is a random variable. Therefore, the indirect utility of brand j for consumer i on purchase occasion t (at the consideration stage), $U_{ij,t}$, will be a sum of two random components: the subjective quality belief, $q_{ij,t}$, and the price, $p_{ij,t}$. To ensure a closed form expression for the consideration set probabilities, we assume the distribution of prices of brand j to be Type 1 Extreme Valued with mean $\bar{p}_j$ and variance $s_{p_j}^2$. This implies that the scale and the location parameters of the price distribution of brand j is given by scale parameter $= \dfrac{p}{\sqrt{6}s_{p_j}}$ and location parameter $= \dfrac{\sqrt{6}s_{p_j}e_c}{p} + \overline{p}_j$. (A2.15)

Further, as noted earlier, the consumer's subjective quality beliefs about brand j are normally distributed with mean $w_{ij,t}$ and variance $s^2_{w_{ij,t}}$. The consumer knows the expected value of her subjective quality beliefs $w_{ij,t}$, but does not the true quality q_j. Therefore, the consumer's indirect utility (at the consideration stage) for brand j can be expressed as a sum of the fixed and the random components as

$$U_{ij,t} = qw_{ij,t} - \overline{p}_{\mathrm{j}} + u_{q_{ij,t}} + u_{p_j} \tag{A2.16}$$

In equation (A2.16), $u_{q_{ij,t}}$ is a normal random variable with mean zero and variance $s^2_{w_{ij,t}}$ and u_{p_j} is an EV random variable with mean 0 and variance $s^2_{p_j}$. If the consumer searches for the price of brand j, then she would realize its price at the end of the consideration stage; in other words, u_{p_j} would be revealed to her. But the uncertainty in the quality of brand j, $u_{q_{ij,t}}$, will not be revealed to her even at the end of the consideration stage. Therefore, the consumer will only be interested in the expected value of the quality of brand j, $w_{ij,t}$, and not in the random component of the quality $u_{q_{ij,t}}$. Said differently, given the inherent uncertainty in brand qualities, the consumer makes her consideration set as well as brand choice decision so as to maximize her expected surplus (indirect utility).

Note that given these assumptions, the indirect utility of brand j at the consideration stage is an EV random variable with mean $qw_{ij,t} - \overline{p}_j$ and variance $s^2_{p_j}$. This implies that the scale and the location parameters of the price distribution of brand j is given by

scale parameter $= \dfrac{p}{\sqrt{6}s_{p_j}}$ and location parameter $= \dfrac{\sqrt{6}s_{p_j}e_c}{p} - qw_{ij,t} + \overline{p}_j$. (A2.17)

Using mathematical notation, at the consideration stage, we have the following distribution for the consumer i's surplus associated with brand j at time t:

$$U_{ij,t} \sim EV\left(\frac{p}{\sqrt{6}s_{p_j}}, \frac{\sqrt{6}s_{p_j}e_c}{p} - qw_{ij,t} + \overline{p}_j\right). \tag{A2.18}$$

We assume that to search the posted price of brand j on purchase occasion t, the consumer i has to incur a certain search cost $C_{ij,t}$. We further assume that the consumer adopts a fixed sample search strategy for searching the prices. Thus, her optimal consideration set, $\{k\}_{it}$, is the set of all brands that maximizes the difference of the expected value of the utility maximizing brand at the consideration set and the total search costs for searching the prices of all the brands in that set. This can be written as follows:

$$\{k\}_{ij,t} = \arg\max_{\{h\}} E\max\left(\{U_{ij,t}\}_{j\in\{h\}}\right) - \sum_{j\in\{h\}} C_{ij,t} \tag{A2.19}$$

If we assume the variances of the prices of all the brands $s^2_{p_j}$ to be the same (that is, $s^2_{p_j} \equiv s^2_p$ for all brands j), then we can get a closed form expression for $E\max\left(\{u_{ij,t}\}_{j\in\{h\}}\right)$. In that case, $\max\left(\{u_{ij,t}\}_{j\in\{h\}}\right)$ will also have an EV distribution with the scale parameter, a, and location parameter, b, given by

$$a = \frac{p}{\sqrt{6}s_p} \quad \text{and} \quad b = \frac{\sqrt{6}s_p}{p} \ln\left[\sum_{j\in\{h\}} \exp\left(\frac{p}{\sqrt{6}s_p}\left(-qw_{ij,t} + \overline{p}_j\right)\right)\right] + \frac{\sqrt{6}s_p e_c}{p}. \tag{A2.20}$$

Here we have exploited the fact that the maximum of N EV random variables x_j, $j = 1 \ldots N$ with the same scale parameter a and location parameters b_j, $j = 1 \ldots N$ is also distributed EV with the scale parameter a and location parameter b that is related to the scale parameter a and location parameters b_j's (Johnson and Kotz, 1974) as follows:

$$U \equiv \max\left\{u_j \mid u_j \sim EV\left(a, b_j\right)\right\}_{j=1}^{N} \sim \prod_{j=1}^{N} EV\left(a, b_j\right) \tag{A2.21}$$

$$\sim EV\left(a, -\frac{1}{a} \ln\left[\sum_{j=1}^{1} \exp\left(-ab_j\right)\right]\right)$$

Therefore, the expected maximum utility by selecting the utility-maximizing brand from the set $\{h\}$ is given by

$$E \max\left(\left\{U_{ij,t}\right\}_{j\in\{h\}}\right) = \frac{\sqrt{6}s_p}{p} \ln\left(\sum_{j\in\{h\}} \exp\left(\frac{p}{\sqrt{6}s_p}\left(qw_{ij,t} - \overline{p}_j\right)\right)\right). \tag{A2.22}$$

Thus, the consumer's optimal consideration set, $\{k\}_{it}$, is given by

$$\{k\}_{i,t} = \arg\max_{\{h\}} \frac{\sqrt{6}s_p}{p} \ln\left(\sum_{j\in\{h\}} \exp\left(\frac{p}{\sqrt{6}s_p}\left(qw_{ij,t} - \overline{p}_j\right)\right)\right) - \sum_{j\in\{h\}} C_{ij,t}. \tag{A2.23}$$

2.3 Derivation of the Consideration Probabilities for the Proposed Model – Analyst's Perspective:

In equation (A2.23), the consumer i knows deterministically the values of $w_{ij,t}$ for all brands j since she has observed the realizations of the quality cues $\hat{l}_{ij,t-1}$. Hence, the consumer knows precisely the optimal consideration set $\{k\}_{it}$. However, the analyst can only make a probabilistic estimate of the mean quality belief of brand j at time t (as represented by $\tilde{w}_{ij,t}$). From the analyst's perspective, $\tilde{w}_{ij,t}$ is a sum of truncated normal random variables as given in equation (A2.12). It follows that the analyst can only make a probabilistic statement about the consumer's optimal consideration set. Specifically, to the analyst, the probability that any set $\{k\}_{it}$ is the optimal consideration set for consumer i on purchase occasion t is given by

$$\Pr\left[\{k\}_{it} = \text{Con. Set}\right] = \Pr\left[\{k\}_{it} = \arg\max_{\{j\}} \frac{\sqrt{6}s_p}{p} \ln\left(\sum_{l\in\{j\}} \exp\left(\frac{p}{\sqrt{6}s_p}\left(q\tilde{w}_{il,t} - \overline{p}_l\right)\right)\right) - \sum_{l\in\{j\}} C_{il,t}\right] \tag{A2.24}$$

where

$$\tilde{w}_{ij,t} = \tilde{w}_{ij,t-1} + \tilde{w}_{ij,t}^{N}, \tag{A2.25}$$

$$\tilde{w}_{ij,t-1} = \sum_{s=1}^{t-1} \underset{\{R_{i,s}\}_{s=1}^{t-1}}{\text{Truncated}}\left[N\left(0, \frac{d_{ij,s-1}^{1}\left(1 + a_{ij,s-1}^{-1}\right)}{\left(a_{ij,s-1} + d_{ij,s-1}^{1}\right)^2}\right)\right], \text{ and,} \tag{A2.26}$$

$$\tilde{w}_{ij,t}^{N} \sim N\left(0, \frac{d_{ij,t-1}^{1}\left(1 + a_{ij,t-1}^{-1}\right)}{\left(a_{ij,t-1} + d_{ij,t-1}^{1}\right)^2}\right). \tag{A2.27}$$

Note that on the random variables $\{\tilde{w}_{ij,t-1}\}_{j\in\{y\}}$ are subject to set of restrictions $\{R_{i,s}\}_{s=1}^{t-1}$ because of the sequence of brand choices made by the consumer till the purchase occasion t-1. Further note that the random variables $\{\tilde{w}^{N}{}_{ij,t-1}\}_{j\in\{y\}}$ – that correspond to consumption *after* the purchase in period t-1 – "do not have any restrictions on their state space" as a result of the choices made till t-1.

2.4 Consumer's Optimal Brand Choice for the Proposed Model:

At the brand choice stage, the consumer knows the actual posted prices of all the brands in her optimal consideration set $\{k\}_{it}$. Note that the qualities of the brands in her consideration set still remain unknown (i.e. random variables). The consumer purchases the optimal brand n_{it} that gives the highest *expected* indirect utility (consumer surplus) among all the brands in consideration set $\{k\}_{it}$. Now, the expected utility for any brand j in the optimal consideration set (for the consumer i at time t) will be

$$E(u_{ij,t}) = qw_{ij,t} - p_{ij,t} \,. \tag{A2.28}$$

Note that in equation (A2.28), $p_{ij,t}$ refers to the actual posted price that the consumer, having engaged in price search, now knows for all the brands in her consideration set. Thus, the consumer's optimal brand on purchase occasion t will be

$$n_{it} \;=\; \underset{j\in\{k\}_{it}}{\arg\max}\ \left[qw_{ij,t} - p_{ij,t}\right] \tag{A2.29}$$

2.5 Derivation of the Choice Probabilities given the Optimal Consideration Set for the Proposed Model– Analyst's Perspective:

While making her optimal brand choice decision from the optimal consideration set, the consumer i knows deterministically the values of $w_{ij,t}$ for all brands j. So, the optimal brand n_{it} is known deterministically. But again, the analyst can only make probabilistic estimate of the mean quality belief of brand j at time t (as represented by $\tilde{w}_{ij,t}$). From the analyst's perspective, $\tilde{w}_{ij,t}$ is a sum of truncated normal random variables as given in equation (A2.13). Therefore, for the analyst, the probability that any brand r is the optimal brand (given the optimal consideration set $\{k\}_{it}$) at time t will be

$$\Pr\left(r = n_{it} \middle| \{k\}_{it}\right) = \Pr\left(r = n_{it} = \underset{j\in\{k\}_{it}}{\arg\max}\ \ q\tilde{w}_{ij,t-1} + q\tilde{w}^{N}_{ij,t} - p_{ij,t}\right) \tag{A2.30}$$

In equation (A2.30), $\tilde{w}_{ij,t-1}$, and $\tilde{w}^{N}_{ij,t}$ are as defined in equations (A2.27)-(A2.28). Note that choice probabilities in equation (A2.31) are calculated conditional on the fact that $\{k\}_{it}$ is the optimal consideration set. Therefore, while computing the choice probabilities in equation (A2.30), there will be an additional set of restrictions imposed on $\{\tilde{w}_{ij,t}\}_{h\in\{y\}}$ because of the fact that that $\{k\}_{it}$ was the optimal consideration set at time t. These additional restrictions are given by:

$$R_{i1,t} \;\equiv\; \left|\left\{\tilde{w}_{ih,t} = \tilde{w}_{ih,t-1} + \tilde{w}^{N}_{ih,t-1}\right\}_{h\in\{y\}}\right| \text{ such that} \tag{A2.31}$$

$$\{k\}_{i,t} = \underset{\{j\}\in\{y\}}{\arg\max}\ \frac{\sqrt{6}s_{p}}{p}\ \ln\left(\sum_{h\in\{j\}}\exp\left(\frac{p}{\sqrt{6}s_{p}}\left(q\tilde{w}_{ih,t} - \bar{p}_{h}\right)\right)\right) - \sum_{h\in\{j\}}C_{ih,t}\Bigg]$$

3. ESTIMATION ISSUES

3.1 Computation of Consideration and Choice Probabilities – Restrictions on $\{w_{ij,t-1}\}_{j\in\{y\}}$:

The probability that consumer *i* selects brand *m* on purchase occasion *t* is given by

$$\Pr(n_{i,t}=m) \;=\; \sum_{\{g\}\in\{y\}} \Pr(\{k\}_{i,t}=\{g\}) \times \Pr(n_{i,t}=m \mid \{k\}_{i,t}=\{g\}) \qquad \text{(A3.1)}$$

where $\Pr(\{k\}_{i,t}=\{g\})$ is the probability of $\{g\}$ being the optimal consideration set and is given by equation (A2.24) and $\Pr(n_{i,t}=m \mid \{k\}_{i,t}=\{g\})$ is the probability that m is the utility-maximizing brand among all the brands included in the set $\{g\}$.

As noted earlier, while calculating the consideration set probabilities in equation (A2.24) or brand choice probabilities in equation (A2.30), we need to have a set of restrictions on the random variables $\{\tilde{w}_{ij,t-1}\}_{j\in\{y\}}$. Here $\{y\}$ denotes the universal set of brands i.e. all the brands in the product category. These restrictions are needed because of the choice made by the consumer on purchase occasion t-1. We will represent these restrictions on the space of $\{\tilde{w}_{ij,t-1}\}_{j\in\{y\}}$ as $R_{i,t-1}$. We discuss these restrictions on $\{\tilde{w}_{ij,t-1}\}_{j\in\{y\}}$ below.

Consider the case when brand n_{it-1} was bought at time t-1. We will have two sets of restrictions on the expected utilities of all brands at time t-1:

- First, that the brand n_{it-1} was the optimal brand chosen from the optimal consideration set at time t-1; and,
- Second that the optimal consideration set contained the brand n_{it-1}.

These two sets of restrictions can be formalized as follows.

RESTRICTION 1: Brand n_{it-1} was chosen from the optimal consideration set $\{k\}_{i,t-1}$ at time t-1 *given that* $\{k\}_{i,t-1}$ contained brand n_{it-1}:

Define j as any brand in the optimal consideration set $\{k\}_{i,t-1}$ at time t-1. The first set of restrictions, $R_{1,i,t-1}$, on $\{\tilde{w}_{ij,t-1}\}_{h\in\{y\}}$ will be:

$$R_{1,i,t-1} \equiv \left[\{\tilde{w}_{ij,t-1}\}_{j\in\{y\}} \middle| q\tilde{w}_{in_{i,t-1},t-1} - p_{in_{i,t-1},t-1} \geq q\tilde{w}_{ij,t-1} - p_{ij,t-1} \;\; \forall n_{i,t-1},\, j\in\{k\}_{i,t-1}\right] \qquad \text{(A3.2)}$$

RESTRICTION 2: The optimal consideration set $\{k\}_{i,t-1}$ was chosen that contained brand n_{it-1} during the purchase occasion at time t-1:

Define $\{g\}_{n_{it-1}}$ as the set of all possible subsets of brands that contain brand n_{it-1} in them. The second set of restrictions, $R_{2,i,t-1}$, on the random variables $\{\tilde{w}_{ih,t-1}\}_{h\in\{y\}}$ is that the optimal consideration set at time t-1, $\{k\}_{i,t-1}$, is a subset of $\{g\}_{n_{i,t-1}}$.

This can be represented mathematically as:

$$R_{2,i,t-1} \;\equiv\; \left[\{\tilde{w}_{ih,t-1}\}_{h\in\{y\}} \middle| \{k\}_{i,t-1} \in \{g\}_{n_{i,t-1}}, \;\; \forall\{k\}_{i,t-1} \;\; \text{such that} \right. \qquad \text{(A3.3)}$$

$$\left. \{k\}_{i,t-1} = \arg\max_{\{j\}\in\{y\}} \frac{\sqrt{6}s_p}{p} \ln\left(\sum_{h\in\{j\}} \exp\left(\frac{p}{\sqrt{6}s_p}\left(q\tilde{w}_{ih,t-1} - \bar{p}_h\right)\right)\right) - \sum_{h\in\{j\}} C_{ih,t-1}\right]$$

We can summarize these two sets of restrictions into one set as:

$$R_{i,t-1} \equiv \left[\{\tilde{w}_{ih,t-1}\}_{h\in\{y\}} \middle| \{\tilde{w}_{ih,t-1}\}_{h\in\{y\}} \in R_{1,i,t-1}, R_{2,i,t-1}\right] \tag{A3.4}$$

In order to calculate the consideration probabilities in equation (A2.24) we need to simulate the random variables $\{\tilde{w}_{ij,t-1}\}_{j\in\{y\}}$ and $\{\tilde{w}^N_{ij,t}\}_{j\in\{y\}}$. The random variables $\{\tilde{w}^N_{ij,t}\}_{j\in\{y\}}$ are normal random variables with no restrictions on their state space. So, it is easy to simulate them. On the other hand $\{\tilde{w}_{ij,t-1}\}_{j\in\{y\}}$ are sums of truncated normal random variables as given in (A2.26) that have to satisfy the set of restrictions given by $\{R_{i,s}\}_{s=1}^{t-1}$. Finally while calculating the choice probabilities conditional on a given consideration set, additional restrictions are imposed on $\{\tilde{w}_{ij,t-1} + \tilde{w}^N{}_{ij,t-1}\}_{j\in\{y\}}$ as given in (A2.31).

3.2 Details of the Estimation Procedure for the Proposed Model:

We have to estimate 11 parameters:

(i) Mean Quality sensitivity parameter, $\bar{q}$;

(ii) Variance of the quality sensitivity across the population, s_q^2.

(iii) 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, a_0;

(iv) Base line search cost per brand for households with no full time house maker, in the absence of prior store visits, marketing activities (features and displays) and with income per member of 10,000 dollars, C_0;

(v) Effect of the presence of displays on the search costs of the household, C_1;

(vi) Effect of the presence of feature advertisements on search costs of the household, C_2;

(vii) Effect of store familiarity on search costs of the household, C_3;

(viii) Effect of the day of the week (when the brand was purchased) on the search costs of the household, C_4;

(ix) Effect of the presence of a full time house maker on the search costs of the household, C_5;

(x) Effect of increase of the income per member of the household by 1,000 dollars on the search costs of the household, C_6;

(xi) Effect of the presence of display for a brand in the last time period on the search cost for the brand, C_7;

We have used the Method of Simulated Moments (MSM) to estimate these 11 parameters. We have used 15 instruments:

- Price, $P_{ij,t}$;
- Square of the price, $P_{ij,t}{}^2$;
- Cube of the price, $P_{ij,t}{}^3$;
- Quadruple of the price, $P_{ij,t}{}^4$;
- Lagged price, $P_{ij,t-1}$;
- Square of the lagged price, $P_{ij,t-1}{}^2$;
- Cube of the lagged price, $P_{ij,t-1}{}^3$;

- Quadruple of the lagged price, $P_{ij,t-1}{}^4$;
- Past purchase, $d_{ij,t-1}$;
- Purchase two time periods back $d_{ij,t-2}$
- Purchase three time periods back $d_{ij,t-3}$
- Whether the brand was feature advertised, $d^2_{ij,t}$;
- Whether the brand was on display, $d^1_{ij,t}$;
- Whether the brand was feature advertised in the last purchase occasion, $d^2_{ij,t-1}$; and,
- Whether the brand was on display during the last purchase occasion, $d^1_{ij,t-1}$.

We have generated ten sets of the random variables $\{w^N_{ij,t}\}_{j\in\{y\}}$ and $\{w_{ij,t-1}\}_{j\in\{y\}}$ to calculate every probability expression in (3.1). This ensures that we get a consistent estimate for our set of parameters.

4. MODEL II – QUALITY LEARNING WITH NO CONSIDERATION STAGE

4.1 Additional Notations:

SYMBOL	EXPLANATION
$y^{d}_{ij,t}$	Quality signal associated with displays for brand j received at time t by the household i. This signal is assumed to be “noisy” so that $y^{d}_{ij,t} \sim N\left(q_j, s^2_{dn}\right)$. Note that the consumer observes the realized value of this signal; however, from the analyst’s perspective, it is an unobservable.
$n^{d}_{ij,t}$	“Noise” in the quality signal received by the household i about brand j if that brand is displayed at time t. Thus, $y^{d}_{ij,t} = q_j + n^{d}_{ij,t}$.
s^2_{dn}	Variance of the quality signal associated with displays. It is assumed to be the same across all households and across all brands and across all purchase occasions.
k_d	$k_d = s^2_h / s^2_{dn}$. This is the ratio of variance of the “noise” terms in the quality signals associated with consumption experience and displays. It measures the relative informativeness (about the “true” brand quality) of displays and consumption. Thus, $k_d < 1$ would imply that consumption experience is more informative than displays and vice versa.
$y^{f}_{ij,t}$	Quality signal associated with feature advertisement for brand j received at time t by the household i. This signal is assumed to be “noisy” so that $y^{f}_{ij,t} \sim N\left(q_j, s^2_{fn}\right)$. Note that the household observes the realized value of this signal; however, from the analyst’s perspective, it is an unobservable.
$n^{f}_{ij,t}$	“Noise” in the quality signal received by the household i about brand j if that brand is feature advertised at time t. Thus, $y^{f}_{ij,t} = q_j + n^{f}_{ij,t}$.
s^2_{fn}	Variance of the quality signal associated with feature advertising. It is assumed to be the same across all households and across all brands and across all purchase occasions.
k_f	$k_f = s^2_h / s^2_{fn}$. This is the ratio of variance of the “noise” terms in the quality signals associated with consumption experience and feature advertisements. It measures the relative informativeness (about the “true” brand quality) of feature advertisement and consumption. Thus, $k_f < 1$ would imply that consumption experience is more informative than feature advertising and vice versa.
$d^{1}_{ij,t}$	Indicator variable such that $d^{1}_{ij,t}$ = 1 if brand j is displayed on household i’s tth purchase occasion.
$d^{2}_{ij,t}$	Indicator variable such that $d^{2}_{ij,t}$ = 1 if brand j is feature advertised on household i’s tth purchase occasion.

4.2 Evolution of the Quality Perception for Model II:

Similar to the proposed model, we assume that the perceived quality of brand j on purchase occasion t for the consumer i, $q_{ij,t}$, is a normal random variable with mean $w_{ij,t}$ and variance $s^2_{w_{ij,t}}$. We assume consumer can learn about the "true" qualities of brand j, q_j, through (i) Consumption experience in the previous time period; (ii) Displays in the present time period; and, (iii) Feature advertisement in the present time period. Like in Model 1, we assume these mechanisms only provide "noisy" signals of the "true" quality. We operationalize consumer learning about brand qualities through "noisy" consumption, display and feature advertising signals as follows:

Let $y^d_{ij,t}$, $y^f_{ij,t}$ and $l_{ij,t-1}$ denote the quality cues associated with displays, feature advertising message and consumption experience respectively received by consumer i about brand j specified as follows:

$$y^d_{ij,t} = q_j + n^d_{ij,t}, \tag{A4.1}$$

$$y^f_{ij,t} = q_j + n^f_{ij,t} \tag{A4.2}$$

$$l_{ij,t-1} = q_j + h_{ij,t-1}. \tag{A4.3}$$

In equations (A4.1) to (A4.2), q_j denotes the "true" quality of brand j. Further, $n^d_{ij,t}$ denotes the noise associated with the feature advertising message, $n^f_{ij,t}$ denotes the noise associated with the feature advertising message and $h_{ij,t-1}$ denotes the noise associated with the consumption signal. We assume that $h_{ij,t-1}$, $n^d_{ij,t}$ and $n^f_{ij,t}$ are *i.i.d* across all consumers, across all brands and across all purchase occasions. We further assume that $h_{ij,t-1}$, $n^d_{ij,t}$ and $n^f_{ij,t}$ are uncorrelated with each other.

To exploit the self-conjugacy of the normal density, we assume that $h_{ij,t-1}$, $n^d_{ij,t}$ and $n^f_{ij,t}$ are normal random variables with zero mean and variances equal to s^2_h, s^2_{dn} and s^2_{fn} respectively, i.e.,

$$h_{ij,t-1} \sim N\left(0, s^2_h\right),\ n^d_{ij,t} \sim N\left(0, s^2_{dn}\right) \text{ and } n^f_{ij,t} \sim N\left(0, s^2_{fn}\right). \tag{A4.4}$$

Thus, s^2_h is a measure of the non-informativeness of the consumption experience, s^2_{dn} measures the non-informativeness of the displays and s^2_{fn} measures the non-informativeness of the advertisements. Thus, if either $s^2_h = 0$ or $s^2_{fn} = 0$ or $s^2_{dn} = 0$, the consumer will get to learn the "true" brand quality immediately.

Consider consumer i who might receive consumption and/or feature advertising and/or display about the quality of brand j between t-1 and t purchase occasions. Let $d^1_{ij,t-1}$ be the indicator variable such that

$d^1_{ij,t-1}$	=	1	if consumer i receives a display message $y^d_{ij,t}$ for brand j at the purchase occasion t; and,
	=	0	otherwise.

Similarly, define indicator variable $d^2_{ij,t}$ such that

$d^2_{ij,t}$	=	1	if consumer i receives an advertising message $y^f_{ij,t}$ for brand j at the purchase occasion t; and,
	=	0	otherwise.

The consumer has prior subjective beliefs about the quality of brand j, based on her purchase history at time t-1, $\mathrm{H}_i(t-1)$, before she observes the consumption signal, $l_{ij,t-1}$, the display signal $y^{d}{}_{ij,t}$ and the feature advertising signal, $y^{f}{}_{ij,t}$. Let the prior beliefs of the consumer i be denoted by

$$q_{j,t} \mid \mathrm{H}_i(t-1) \sim N\left(w_{ij,t-1}, s^{2}_{w_{ij,t-1}}\right). \tag{A4.5}$$

After observing the quality cues $\hat{y}^{d}{}_{ij,t}$, $\hat{y}^{f}{}_{ij,t}$ and $\hat{l}_{ij,t-1}$, the consumer updates her subjective beliefs about the quality of brand j in a Bayesian fashion. Represent her posterior beliefs, after observing the quality cues $\hat{y}^{d}{}_{ij,t}$, $\hat{y}^{f}{}_{ij,t}$ and $\hat{l}_{ij,t-1}$, by

$$q_{j,t} \mid \mathrm{H}_i(t) \sim N\left(w_{ij,t}, s^{2}_{w_{ij,t}}\right). \tag{A4.6}$$

Then, because of the self-conjugacy of the normal density, the mean and variance of the posterior beliefs are related to the mean and variance of the prior beliefs as (DeGroot, 1970):

$$w_{ij,t} = \frac{\dfrac{w_{ij,t-1}}{s^{2}_{w_{ij,t-1}}} + d_{ij,t-1}\dfrac{\hat{l}_{ij,t-1}}{s^{2}_{h}} + d^{1}_{ij,t}\dfrac{\hat{y}^{d}{}_{ij,t}}{s^{2}_{dn}} + d^{2}_{ij,t}\dfrac{\hat{y}^{f}{}_{ij,t}}{s^{2}_{fn}}}{\dfrac{1}{s^{2}_{w_{ij,t-1}}} + d_{ij,t-1}\dfrac{1}{s^{2}_{h}} + d^{1}_{ij,t}\dfrac{1}{s^{2}_{dn}} + d^{2}_{ij,t}\dfrac{1}{s^{2}_{fn}}}, \text{ and,} \tag{A4.7}$$

$$\frac{1}{s^{2}_{w_{ij,t}}} = \frac{1}{s^{2}_{w_{ij,t-1}}} + d_{ij,t-1}\frac{1}{s^{2}_{h}} + d^{1}_{ij,t}\frac{1}{s^{2}_{dn}} + d^{2}_{ij,t}\frac{1}{s^{2}_{fn}} \tag{A4.8}$$

Define

$$a_{ij,t-1} = s^{2}_{h}\big/s^{2}_{w_{ij,t-1}}, \; k_d = s^{2}_{h}\big/s^{2}_{dn} \text{ and } k_f = s^{2}_{h}\big/s^{2}_{fn}. \tag{A4.9}$$

The, we can rewrite equations (A4.7) and (A4.8) as

$$w_{ij,t} = \frac{w_{ij,t-1}a_{ij,t-1} + d_{ij,t-1}\hat{l}_{ij,t-1} + k_d\, d^{1}_{ij,t}\hat{y}^{d}{}_{ij,t} + k_f\, d^{2}_{ij,t}\hat{y}^{f}{}_{ij,t}}{a_{ij,t-1} + d_{ij,t-1} + k_d\, d^{1}_{ij,t} + k_f\, d^{2}_{ij,t}}, \text{ and,} \tag{A4.10}$$

$$a_{ij,t} = a_{ij,t-1} + d_{ij,t-1} + k_d\, d^{1}_{ij,t} + k_f\, d^{2}_{ij,t} \tag{A4.11}$$

Equation (A4.10) shows how the mean of the subjective quality beliefs about brand j evolves when the consumer i receives the consumption, display and feature advertising signals. The consumer observes the signals $\hat{y}^{d}{}_{ij,t}$, $\hat{y}^{f}{}_{ij,t}$ and $\hat{l}_{ij,t-1}$, Hence, she knows deterministically the values of $w_{ij,t}$ for all brands j at any time t. But the analyst does not observe the realization of the random variables $\hat{y}^{d}{}_{ij,t}$, $\hat{y}^{f}{}_{ij,t}$ and $\hat{l}_{ij,t-1}$. The analyst can only make a probabilistic estimate of the consumer i's mean quality belief of brand j at time t (which we denote by $\tilde{w}_{ij,t}$). From the analyst's perspective, $\tilde{w}_{ij,t}$ is a random variable whose characterization can be obtained by substituting $l_{ij,t-1} = q_{ij,t-1} + h_{ij,t-1}$, $y^{d}{}_{ij,t} = q_{ij,t-1} + n^{d}{}_{ij,t}$ and $y^{f}{}_{ij,t} = q_{ij,t-1} + n^{f}{}_{ij,t}$ in equation (A4.10) to get

$$\tilde{w}_{ij,t} = \tilde{w}_{ij,t-1} + N\left(0, \frac{d_{ij,t-1}\left(1 + a^{-1}_{ij,t-1}\right) + k_d\, d^{1}_{ij,t}\left(1 + k_d\, a^{-1}_{ij,t-1}\right) + k_f\, d^{2}_{ij,t}\left(1 + k_f\, a^{-1}_{ij,t-1}\right)}{\left(a_{ij,t-1} + d_{ij,t-1} + k_d d^{1}_{ij,t} + k_f d^{2}_{ij,t}\right)^{2}}\right) \tag{A4.12}$$

We assume the means of the quality beliefs for all brands and for all the consumers at the beginning of their consumption history are zero. Also, we assume that the value of the precision of the subjective quality beliefs of all the brands for all the consumers at the beginning of their consumption history is a_0.

In other words, $\tilde{w}_{ij,0} = 0$ and $a_{ij,0} = a_0$ for all brands j and for all consumers i. We can then recursively derive the expressions for $\tilde{w}_{ij,t}$ and $a_{ij,t}$ as a sum of all previous advertising, display and consumption signals as

$$\tilde{w}_{ij,t} = N\left(0, \frac{d_{ij,t-1}\left(1+a_{ij,t-1}^{-1}\right)+d_{ij,t}^{1}k_d\left(1+k_d\, a_{ij,t-1}^{-1}\right)+d_{ij,t}^{2}k\left(1+k_f\, a_{ij,t-1}^{-1}\right)}{\left(a_{ij,t-1}+d_{ij,t-1}+k_d d_{ij,t}^{1}+k_f d_{ij,t}^{2}\right)^2}\right)+ \sum_{s=1}^{t-1}\underset{\{R_{i,s}\}_{s=1}^{t-1}}{\text{Truncated}}N\left(0, \frac{d_{ij,s-1}\left(1+a_{ij,s-1}^{-1}\right)+d_{ij,s}^{1}k_d\left(1+k_d\, a_{ij,s-1}^{-1}\right)+d_{ij,s}^{2}k\left(1+k_f\, a_{ij,s-1}^{-1}\right)}{\left(a_{ij,s-1}+d_{ij,s-1}+k_d d_{ij,s}^{1}+k_f d_{ij,s}^{2}\right)^2}\right) \quad \text{(A4.13)}$$

$$a_{ij,t} = a_0 + \sum_{s=1}^{t}\left(d_{ij,s-1}+k_d\, d_{ij,s}^{1}+k_f\, d_{ij,s}^{2}\right) \quad \text{(A4.14)}$$

4.3 Derivation of the Choice Probabilities:

While making her optimal brand choice decision from the universal set, the consumer i knows deterministically the values of $w_{ij,t}$ for all brands j. So, the optimal brand n_{it} is deterministically known. But again, the analyst does not observe the values of $w_{ij,t}$ and can only make a probabilistic estimate of it (as given by $\tilde{w}_{ij,t}$). For the analyst, $\tilde{w}_{ij,t}$ is the sum of truncated normal random variables, which is given by equation (A4.13). Therefore, for the analyst, the probability that any brand k is the optimal brand at time t will be

$$\Pr\left(k = n_{it}\right) = \Pr\left(k = n_{it} = \arg\max_{j}\ q\tilde{w}_{ij,t-1} + q\tilde{w}_{ij,t}^{N} - p_{ij,t} \middle| \{\tilde{w}_{ij,t-1}\}_{j\in\{y\}} \in R_{i,t-1}\right) \quad \text{(A4.15)}$$

In equation (A4.15), $\tilde{w}_{ij,t-1}$, and $\tilde{w}_{ij,t}^{N}$ are defined as
where

$$\tilde{w}_{ij,t} = \tilde{w}_{ij,t-1} + \tilde{w}_{ij,t}^{N}, \quad \text{(A4.16)}$$

$$\tilde{w}_{ij,t-1} = \sum_{s=1}^{t-1}\underset{\{R_{is}\}_{s=1}^{t-1}}{\text{Truncated}}N\left(0, \frac{d_{ij,s-1}\left(1+a_{ij,s-1}^{-1}\right)+d_{ij,s}^{1}k_d\left(1+k_d\, a_{ij,s-1}^{-1}\right)+d_{ij,s}^{2}k\left(1+k_f\, a_{ij,s-1}^{-1}\right)}{\left(a_{ij,s-1}+d_{ij,s-1}+k_d d_{ij,s}^{1}+k_f d_{ij,s}^{2}\right)^2}\right) \quad \text{(A4.17)}$$

$$\tilde{w}^{N}_{ij,t} \sim N\left(0, \frac{d_{ij,t-1}\left(1+a_{ij,t-1}^{-1}\right)+d_{ij,t}^{1}k_d\left(1+k_d\, a_{ij,t-1}^{-1}\right)+d_{ij,t}^{2}k\left(1+k_f\, a_{ij,t-1}^{-1}\right)}{\left(a_{ij,t-1}+d_{ij,t-1}+k_d d_{ij,t}^{1}+k_f d_{ij,t}^{2}\right)^2}\right). \quad \text{(A4.18)}$$

Note that while calculating the probabilities in (A4.15), we need to set of restrictions on the random variables $\left\{\tilde{w}_{ij,t-1}\right\}_{j\in\{y\}}$ (on the mean of the quality beliefs of all the brands at time t-1) because of the sequence of brand choices made by the consumer till the purchase occasion t-1. We will represent these restrictions on the space of $\left\{\tilde{w}_{ij,t-1}\right\}_{j\in\{y\}}$ as $R_{i,t-1}$. Consider the case when brand n_{it-1} was bought at time t-1. We will have one set of restrictions on the expected utilities of all brands at time t-1:

- The brand n_{it-1} was the optimal brand chosen from the universal set at time t-1;

This restriction can be formalized as follows.

<u>RESTRICTION</u>: Brand n_{it-1} was chosen from the universal set at time t-1*:*

Define j as any brand in the universal set at time t-1. The set of restrictions, $R_{i,t-1}$, on $\left\{\tilde{w}_{ij,t-1}\right\}_{j\in\{y\}}$ will be:

$$R_{i,t-1} \equiv \left[\left\{ \tilde{w}_{ij,t-1} \right\}_{j \in \{y\}} \middle| q\tilde{w}_{in_{i,t-1},t-1} - p_{in_{i,t-1},t-1} \geq q\tilde{w}_{ij,t-1} - p_{ij,t-1} \quad \forall n_{i,t-1},\, j \in \{y\} \right] \tag{A4.19}$$

In order to calculate the probabilities in equations (A4.15), we need to simulate the random variables $\left\{\tilde{w}_{ij,t-1}\right\}_{j\in\{y\}}$ and $\left\{\tilde{w}^{N}_{ij,t}\right\}_{j\in\{y\}}$. The random variables $\left\{\tilde{w}^{N}_{ij,t}\right\}_{j\in\{y\}}$ are normal random variables with no restrictions on their state space. So, it is easy to simulate them. On the other hand $\left\{\tilde{w}_{ij,t-1}\right\}_{j\in\{y\}}$ are sums of truncated normal random variables as given in (A4.17) that have to satisfy the set of restrictions given by $\left\{R_{i,s}\right\}_{s=1}^{t-1}$.

5. TABLES AND FIGURES

TABLE TA.1: Posterior Quality Beliefs of Brands from the Pre-Estimation Sample[6] for Liquid Detergent Data Set

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

TABLE TA.2: Posterior Quality Beliefs of Brands from the Pre-Estimation Sample[7] for Ketchup Data Set

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

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

TABLE TA.3: Parameter Estimates for Model III for Liquid Detergent Data Set

Parameter	Explanation	MODEL III (Std. Deviation)
U_l	Threshold Utility for Consideration	0.6241 (0.191)
w_{Tide}	Intercept for Tide (w_{Wisk} = 0 for identification)	0.9918 (0.001)
w_{Era}	Intercept for Era	0.7470 (0.002)
w_{Surf}	Intercept for Surf	0.5899 (0.002)
P_{co}	Effect of Price on Probability of Consideration	2.0204 (1.316)
A_{co}	Effect of Feature Advertisement on Probability of Consideration	1.0589 (0.487)
D_{co}	Effect of Display on Probability of Consideration	1.4722 (0.440)
L_{co}	Effect of Brand Loyalty on Probability of Consideration	2.6065 (0.052)
P_{ch}	Effect of Price on Brand Choice Probability (conditional on consideration)	- 1.8629 (0.003)
A_{ch}	Effect of Feature Ad on Brand Choice Probability (conditional on consideration)	0.4259 (0.401)
D_{ch}	Effect of Display on Brand Choice Probability (conditional on consideration)	0.5437 (0.212)
L_{ch}	Effect of Brand Loyalty on Brand Choice Probability (conditional on consideration)	2.8252 (0.156)

TABLE TA.4: Parameter Estimates for Model IV for Liquid Detergent Data Set

Parameter	Explanation	MODEL IV (Std. Deviation)
l	Cut-off Threshold for Past Purchases	0.2331 (0.0010)
q	Mixing probability for the model in which any brand that is on promotion enters the Consideration Set	0.4106 (0.0700)
g	Mixing probability for the model wherein there is no consideration	0.1727 (0.0401)

Parameter	Explanation	MODEL IV (Std. Deviation)
w_{Tide}	Intercept for Tide (w_{Wisk} = 0 for identification)	0.0110 (0.0021)
w_{Era}	Intercept for Era	- 0.0005 (0.0003)
w_{Surf}	Intercept for Surf	0.0102 (0.0010)
P_{ch}	Effect of Price on Brand Choice Probability (conditional on consideration)	- 0.9022 (0.0322)
A_{ch}	Effect of Feature Ad on Brand Choice Probability (conditional on consideration)	0.7963 (1.3751)
D_{ch}	Effect of Display on Brand Choice Probability (conditional on consideration)	0.8136 (0.5161)
L_{ch}	Effect of Brand Loyalty on Brand Choice Probability (conditional on consideration)	0.5849 (0.0565)

TABLE TA.5: Parameter Estimates for Model V for Liquid Detergent Data Set

Parameter	Explanation	MODEL V (Std. Deviation)
w_{Tide}	Intercept for Tide (w_{Wisk} = 0 for identification)	- 0.0076 (0.0001)
w_{Era}	Intercept for Era	0.0059 (0.0001)
w_{Surf}	Intercept for Surf	- 0.0055 (0.0001)
P_{ch1}	Effect of Price on Brand Choice Probability (conditional on consideration) for Segment 1	- 4.3727 (0.03)
A_{co1}	Effect of Feature Advertisement on Probability of Consideration for Segment 1	0.7013 (0.656)
D_{co1}	Effect of Display on Probability of Consideration for Segment 1	0.4378 (0.241)
L_{co1}	Effect of Brand Loyalty on Probability of Consideration for Segment 1	7.8750 (0.11)
$U_{l\,1}$	Threshold Utility for Consideration for Segment 1	2.2669 (0.67)
P_{ch2}	Effect of Price on Brand Choice Probability (conditional on consideration) for Segment 2	-2.5266 (0.03)
A_{co2}	Effect of Feature Advertisement on Probability of Consideration for Segment 1	0.8290 (0.601)

Parameter	Explanation	MODEL V (Std. Deviation)
D_{co2}	Effect of Display on Probability of Consideration for Segment 1	1.9781 (0.201)
L_{co2}	Effect of Brand Loyalty on Probability of Consideration for Segment 2	6.9023 (0.460)
U_{l2}	Threshold Utility for Consideration for Segment 2	3.1129 (0.730)
l	Relative sizes of the 2 segments [Segment 1's relative size = $e^{\lambda}/(1+ e^{\lambda})$]	1.055 (0.410)

TABLE TA.6: Parameter Estimates for Model IA[8] for Liquid Detergent Data Set

Parameter	Explanation	MODEL IA (Std. Deviation)
$\bar{q}$	Mean value of the willingness-to-pay for quality across the consumers	2.5975 (0.131)
s_q^2	Variance of the willingness-to-pay for quality across the consumers	0.4410 (0.269)
k_f	Ratio of the informativeness (about the true quality) of Feature Advertisement Signal to that of Consumption Signal	0.0629 (0.129)
k_d	Ratio of the informativeness (about the true quality) of Displays to that of Consumption Signal	0.0590 (0.012)
a_0	Inverse of the uncertainty in quality of at the beginning of consumption history; assumed same $\forall i$, $\forall j$	2.9079 (0.732)
C_0	Baseline Search Costs for discovering the posted price of a brand	0.0497 (0.010)
C_1	Effect of Display on Search Costs	-
C_2	Effect of Feature Ad on Search Costs	-
C_3	Effect of Store-Category Familiarity on Search Costs	- 0.0021 (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)

[8]. An alternate specification of Model I when features and displays affect the quality beliefs but not the search costs.

Parameter	Explanation	MODEL IA (Std. Deviation)
C_6	Effect of increase in per capita household income by $1000 on Search Costs	0.0019 (0.0008)
$\bar{w}_{Wisk}$	Mean value of perceived quality of Wisk (across all consumers and across all purchase occasions)	-0.0046 (0.004)
$\bar{w}_{Tide}$	Mean value of perceived quality of Tide	0.0694 (0.010)
$\bar{w}_{Era}$	Mean value of perceived quality of Era	0.0303 (0.008)
$\bar{w}_{Surf}$	Mean value of perceived quality of Surf	0.0376 (0.003)

TABLE TA.7: Result from comparison of MODEL I with MODEL IA

Model Comparison using Singleton Test	Singleton Test Results	
	Estimation Sample	Hold-out Sample
MODEL I against MODEL IA H_0: MODEL IA is the "true" model H_1: MODEL I is the "true" model	$J_S = 3.91$ (c^2 with 1 d.f.) p-value < 0.05	$J_S = 4.15$ (c^2 with 1 d.f.) p-value < 0.05

TABLE TA.8: Inference Results for Hold out sample for Liquid Detergents for Model I and Model IA

	MODEL I	MODEL IA
Hit Rate	71.76%	67.14%
-Log Likelihood	1890	2101

TABLE TA.9: Parameter Estimates for Model I for Ketchup Data Set

Parameter	Explanation	MODEL I (Std. Deviation)
$\bar{q}$	Mean value of quality sensitivity across the consumers	3.2140 (0.25)
s_q^2	Variance of quality sensitivity across the consumers	0.4739 (0.30)
a_0	Inverse of the uncertainty in quality of at the beginning of consumption history; assumed same $\forall i, \forall j$	3.5534 (1.10)
C_0	Baseline Search Costs for discovering the posted price of a brand	0.0284 (0.004)
C_1	Effect of Display on Search Costs	- 0.0105 (0.005)
C_2	Effect of Feature Ad on Search Costs	- 0.0142 (0.007)
C_3	Effect of Store-Category Familiarity on Search Costs	- 0.0001 (0.033)
C_4	Effect on Search Costs if the purchase was made during the weekend	- 0.0002 (0.01)
C_5	Effect of absence of full time homemaker in household on Search Costs	0.0001 (0.02)
C_6	Effect of increase in per capita household income of $1000 on Search Costs	0.0003 (0.01)
$\bar{w}_{Heinz}$	Mean value of the perceived quality of Heinz (across all consumers and across all purchase occasions)	0.1595 (0.006)
$\bar{w}_{Hun's}$	Mean value of the perceived quality of Hunt's	0.0199 (0.002)
$\bar{w}_{DelMonte}$	Mean value of the perceived quality of Del Monte	0.0291 (0.007)
$\bar{w}_{Generic}$	Mean value of the perceived quality of the Generic brand	0.0092 (0.001)

TABLE TA.10: Parameter Estimates for Model IV for Ketchup Set

Parameter	Explanation	MODEL IV (Std. Deviation)
l	Cut-off Threshold for Past Purchases	0.1701 (0.00)
q	Mixing probability for the model in which any brand that is on promotion enters the Consideration Set	0.1455 (0.060)
g	Mixing probability for the model wherein there is no consideration	0.0406 (0.020)
w_{Tide}	Intercept for Hunts (w_{Wisk} = 0 for identification)	-2.1608 (0.0141)
w_{Era}	Intercept for Del Monte	- 2.5186 (0.0140)
w_{Surf}	Intercept for Generic Brand	-2.6157 (0.0141)
P_{ch}	Effect of Price on Brand Choice Probability (conditional on consideration)	- 3.4160 (0.01)
A_{ch}	Effect of Feature Ad on Brand Choice Probability (conditional on consideration)	0.6592 (0.300)
D_{ch}	Effect of Display on Brand Choice Probability (conditional on consideration)	0.4822 (0.01)
L_{ch}	Effect of Brand Loyalty on Brand Choice Probability (conditional on consideration)	1.9385 (0.0173)

TABLE TA.11: Result from comparison of Model I with Model IV for Ketchup

Model Comparison using Singleton Test	Singleton Test Results	
	Estimation Sample	Hold-out Sample
MODEL I against MODEL IV H_0 : MODEL IV is the “true” model H_1 : MODEL I is the “true” model	J_S = 3.01 (c^2 with 1 d.f.) p-value < 0.10	J_S = 3.91 (c^2 with 1 d.f.) p-value <0.05

TABLE TA.12: Hold out sample for Ketchup

	MODEL I	MODEL IV
Hit Rate	74.7%	71.4%
-Log Likelihood	912	965