A STRUCTURAL ESTIMABLE MODEL OF FORGETTING IN MEMORY-BASED CHOICE DECISIONS

Nitin Mehta
Rotman School of Management
University of Toronto
(Email: nmehta@rotman.utoronto.ca)

Surendra Rajiv
Graduate School of Business
National University of Singapore
(Email: rajiv@nus.sng.edu)

Kannan Srinivasan
Graduate School of Industrial Administration
Carnegie Mellon University
(Email: kannans@andrew.cmu.edu)

April 2003

PRELIMINARY DRAFT; PLEASE DO NOT CITE

This work was partly supported by the Cannaught Research Grant (at RSM, University of Toronto) to the first author. The usual disclaimer applies.
ABSTRACT

We propose a structural model to investigate the impact of forgetting on consumer’s brand choice decision in frequently purchased products. Forgetting results in consumers imperfectly recalling their prior brand evaluations while making a purchase decision in the category. We conceptualize the imperfect recall by positing that the consumers recall their prior evaluations with an added noise. The extent of the added noise is modeled by incorporating the results of the experimental work done on forgetting that suggest the extent of imperfect recall would be an increasing and concave function of time. Our model generates interesting analytical insights showing the impact of forgetting on consumer’s brand evaluations and their subsequent purchase behavior. We calibrate our model on the scanner panel data for liquid detergents. We further obtain interesting empirical insights regarding the consumers’ extent of forgetting in the category and on the implications on the extent of learning and the predicted price elasticities.

Key Words: Bayesian Learning, Forgetting, Quality Uncertainty, Structural Model, Econometric Estimation and Memory-Based Choice Decision
1. Introduction

In this paper, we propose a structural estimable model to investigate the impact of forgetting on consumer purchases in frequently purchased products. When making a brand choice in a product category, consumers first construct their evaluations of the qualities of all the alternatives. They do so in at least two different ways (Alba, Hutchinson and Moore, 1991): first, by processing the information of the relevant attributes of the brand that are given on its package; and second, by simply retrieving from their memories the quality evaluations of that brand that they had constructed during their prior purchase in the category. Past research suggests that consumers prefer to retrieve the quality evaluations from their memories if they are time constrained while making their purchase decision (Sonbanmatsu and Fazio 1990), if their purchase decision is made under relatively low involvement conditions (Sonbanmatsu and Fazio 1990, Fazio 1990), and if they have the prior evaluations of all brands stored in their memories as a result of frequent purchases in the category (Beattie and Mitchell 1984, Lichtenstein and Srull 1985, Hastie and Park 1986). Since all these conditions would generally hold true when consumers make purchases in frequently purchased products, it would suggest that in such a context the retrieval of their prior quality evaluations would play an important role in their brand choice.

Forgetting occurs when consumers imperfectly recall their prior overall quality evaluations of the brands while making their purchase decision. This happens for one of the three main reasons (Anderson 1999). First, because memories of the prior quality evaluations weaken with time making them harder to retrieve. Second, because competition from other memories blocks the retrieval of the prior overall evaluations, and third because consumers lose access to the cues that would serve to access the prior overall evaluations at a later point in time. In any case, forgetting results in consumers forming their overall quality evaluations of the brands from a biased information set in their memories.

Experimental work done in the last four decades\(^1\) in social psychology has shown that people’s evaluations based on their prior learning about an object decay over time as a result of forgetting. Despite the overwhelming evidence, prior econometric models of brand choice have typically assumed that learning of the brands’ qualities through consumption persists through time and these quality evaluations of consumers do not decay with time (Erdem and Keane 1996; Mehta, Rajiv and Srinivasan, 2003). Therefore, as a first step towards bridging the gap between the extant econometric models of brand choice and the behavioral literature on memory and

---

\(^1\) See Cook and Flay (1978) for a summary of the literature on decay of people’s evaluations towards an object over time
forgetting, we develop a model of forgetting where consumers imperfectly recall their prior overall evaluations while making their purchase in frequently purchased product categories.

In order to understand how forgetting might impact consumer’s prior brand evaluations, we would first need to address the following set of questions. First, what factors influence the extent of forgetting of consumer’s prior evaluations of brands? And second, how do these overall evaluations depreciate with time as a result of forgetting? The experimental studies in social psychology that have attempted to measure the decay of people’s evaluations over time have shown the following three results. First, people’s evaluations (after they were given additional information about the object in the experiment) converge to their prior evaluations (the evaluations that people had about the object at the beginning of the experiment) over time. Second, the rate of decay of people’s evaluations varies across the individuals and the attitude objects under consideration. And finally, the shape of the decay of people’s evaluations varies across the studies - some studies have shown that people’s evaluations decay in a sigmoid fashion over time fashion (Papageorgis 1963, Watts and McGuire 1964), while the others have shown that the overall evaluations decay in a convex fashion (Collins and Hoyt, 1972); and Ronis et. al., 1977). Therefore, our first objective in this paper is to build a model to forgetting that would (i) explain the two diverse sets of results in the aforementioned studies regarding the shape of the decay of evaluations over time, and (ii) shed an insight into the factors that would influence the extent of decay of the evaluations.

Our second objective is to show that forgetting affects consumer purchase behavior through its impact on the ‘hand of the past’ variables in the consumer’s utility, that is, state dependence and habit persistence (Heckman, 1981; Roy, Chintagunta and Haldar 1996). Doing so would provide a richer insight into why both state dependence and habit persistence vary across consumers and vary across purchase occasions for a given consumer. Even though prior empirical research (Seetharaman, Ainslie and Chintagunta, 1999; and Kim, 2002) has demonstrated the inter-temporal and cross sectional variation of state dependence and habit persistence in a reduced form context, it does not explicitly model the factors that cause these phenomena to occur. We show that incorporating forgetting in Bayesian Learning Models provides an explanation for inter-temporal and cross sectional variation in both habit persistence and state dependence in consumers’ purchase behavior.

1.1 Brief Overview of the Econometric Specification and Main Empirical Findings

The key components of our modeling approach are as follows: the consumer is uncertain about her evaluations of the qualities of all brands in the product category. She forms her overall
evaluation of the quality of a brand by mixing her evaluations from two distinct information sets. The first information set, subsequently referred to as the *positioning set*, consists of the quality evaluation that she infers from the positioning of the brand in the market. This information set is assumed to be time invariant that does not change over the consumer’s purchase history. The second information set, subsequently referred to as the *consumption set*, consists of the evaluation of the brand’s quality that the consumer learns from consumption. While forgetting does not impact the consumer’s evaluations formed from the positioning set, it impacts her evaluations formed from the consumption set. As such, we first detail how the consumer uses her consumption set to form her evaluations of a brand. On a given purchase occasion, the consumer constructs her evaluations based on her consumption set as follows: first notice that she has already constructed her evaluation of each brand based on her past consumption experiences, the last time she shopped in the category. She starts by recalling this – crucially, we suggest that because of forgetting, she recalls this imperfectly. Now clearly some consumption might have taken place since the last time she formed her evaluation (e.g., if she purchased Wisk liquid detergent the last time, her consumption based evaluation today would be different from her last purchase occasion). We suggest that she adds this most current consumption evaluation to her consumption set. This completes our discussion on what constitutes her consumption set at a given purchase occasion. As suggested above, she combines the information in this set with her positioning set to form her overall evaluation of the quality of the brand.

Our structural model yields the following insights about the impact of forgetting on consumer’s overall evaluations of the qualities of the brands. (i) In absence of consumption of a brand, the expected value of consumer’s quality evaluation decays over time in a sigmoid fashion if the information in her consumption set is greater than the information in her positioning set. On the other hand, the consumer’s expected quality evaluation decays in a convex fashion if the information in her consumption set is less than the information in her positioning set. (ii) The greater the information in the consumer’s consumption set, the slower is the decay of her mean quality evaluations with time. (iii) The heterogeneity in both habit persistence and state dependence results from the heterogeneity in the consumers’ willingness to pay for quality and the heterogeneity in their extent of forgetting of prior evaluations. (iv) The habit persistence in consumer’s purchase behavior decreases with the length of the inter-purchase time as a result of forgetting. Further, the greater the information in the consumer’s consumption set, the smaller is the decrease in habit persistence in her purchase behavior over time. (v) The last period’s state dependence in consumer’s utility increases with the length of the inter-purchase time. And the
greater the information in the consumer’s consumption set, the smaller is the increase in state dependence over time.

We calibrate the proposed model using the ERIM data set for liquid detergents. Our empirical analysis yields the following key results: (i) on an average, the consumer’s extent of forgetting of her overall quality evaluations of brands per week is small, but significant. (ii) Ignoring forgetting when it actually occurs leads to the following: (a) an over-estimation of the extent of initial learning and an under-prediction of the extent of learning that can be done that can be done through consumption in the later stages in the consumption history; (b) an over-estimation of the variance in the perceived quality evaluations across the given brands that leads to an under prediction of the consumer’s price elasticity; (c) an over-estimation of the heterogeneity in the price sensitivity across the consumer population.

1.2 Related Literature and Research Contributions

The past relevant literature on forgetting can be classified into two distinct streams: the first stream consists of the experimental studies in forgetting\(^2\) that have attempted to correlate the extent of imperfect recall of information with the time lapsed since the information was encoded in the subject’s memory. In these studies, subjects were typically given constructs like set of words or sentences to memorize at the beginning of the experiment and were later asked to recall them. The extent of imperfect recall at any later point in time was typically measured as the extent of decay of the precision of the recalled information. The main findings of these studies were two fold: (i) the extent of imperfect recall at a later point in time was an increasing and concave function of time lapsed since the information was first encoded; (ii) the extent of imperfect was well captured by a negative exponential, power or a logarithmic function of time. While these studies have only measured the extent of forgetting for constructs like words and sentences, it is reasonable to assume that the same results would hold true in the context of forgetting of prior overall quality evaluations since overall evaluations too are constructs in the memory. In our proposed model, we characterize the extent of forgetting of recalled evaluations as a function of time by incorporating these findings from the past literature. Specifically, we assume the extent of decay of the precision of prior evaluations to be a negative exponential function of the inter-purchase time.

The second stream of literature consists of the theoretical papers that have attempted to model the impact of forgetting on consumer’s consumption decisions. Two notable theoretical\(^2\) See Rubin and Wenzel (1996) for a comprehensive review of the literature in the last one hundred years on characterization of the extent of forgetting as a function of time
papers in this stream are those by Keon (1980) and Mullainathan (2002). Keon (1980) used the findings of the aforementioned studies in experimental psychology to propose a model of forgetting in a reduced form context. He considered the precision of the consumer’s quality evaluation to decay in an exponential fashion in absence of consumption, but assumed the expected value of the overall evaluation remains to be stable over time. He used this model to motivate the fact that in absence of consumption of a brand, a consumer would become uncertain about her overall quality evaluation of that brand and would later switch to it to refresh her memory. Mullainathan (2002) recently proposed a structural model of forgetting to see the impact of the imperfect recall of past predictors of consumer’s future income on her present consumption decision. Mullainathan assumed that consumers either perfectly recall the past predictors or completely forget about their existence and derived the result that the probability of a consumer perfectly recalling a past predictor would be an exponential function of the time lapsed.

The limitations of Keon (1980) are that his model assumes forgetting only impacts the precision and not the expected value of consumer’s overall quality evaluation. Therefore his model would only predict switching to brands that were not consumed in the recent past if the consumers were assumed to be risk taking in their behavior. Unlike Keon (1980), our conceptualization leads to the evolution of consumer’s overall quality evaluations where forgetting impacts both the expected value and the precision of the evaluations. Hence we do not assume consumers to be risk taking in their purchase behavior. Further our result that consumer’s mean evaluations change as a result of forgetting is consistent with the results of past research that has studied the decay of people’s evaluations over time. The limitation of the model proposed by Mullainathan (2002) is that it does not allow for consumers to partially recall the past information – in his conceptualization, consumers either perfectly recall or they do not recall past information at all. On the other hand, we do not assume this simple structure. Our conceptualization leads to consumers recalling distorted versions of their prior quality evaluations as a result of forgetting, which is consistent with the results of the past experimental studies in forgetting.

The rest of the paper is arranged as follows. In Section 2, we develop the model formulation. In Section 3, we discuss the data set used for calibrating the proposed model, report the parameter estimates and discuss the implications of our findings. In Section 4, we conclude and offer directions for future research.
2. MODEL DEVELOPMENT

In this section, we layout the model specification and discuss the main assumptions underlying the proposed formulation. Before developing the model, we provide its conceptual description in section 2.1. In sections 2.2, we lay out the mathematical details underlying the proposed model of forgetting. In section 2.3, we lay out the econometric specification implied by the proposed model.

2.1 Conceptual Description

The basic ingredients of the proposed model are as follows:

(i) At any point in her consumption history, the consumer does not possess enough information to make a precise evaluation about the true quality of any brand in the product category. This is captured by hypothesizing that a consumer has subjective overall quality evaluations (or beliefs) of all the brands.

(ii) A consumer constructs her overall evaluation of each brand’s quality on each purchase occasion. She does so by mixing her quality evaluations of that brand from two different information sets: the first information set, referred to as the *positioning set*, corresponds to the evaluation that the consumer infers about the brand’s quality from its positioning in the market. We assume this evaluation of quality based on the positioning set remains unchanged throughout the consumer’s purchase history\(^3\)^\(^4\).

(iii) The second information set, referred to as the *consumption set*, is the information that the consumer learns about the her true quality evaluation of a brand through consumption. We assume that (i) the consumer cannot learn additionally about quality of the brands while at the store through any conscious information acquisition effort\(^5\). (ii) The consumer updates her evaluation based on consumption experiences every time she purchases and consumes a brand. (iii) The consumption experiences do not fully reveal the brand qualities and only provide a *noisy signal* about them. And finally, (iv) the consumption set decays with time.

---

\(^3\) This assumption is reasonable because: (i) the observation period in our data scans across 111 months within which we can assume that the positioning of the brands did not change. (ii) We would not expect the consumer to forget the evaluations based on the positioning set since the positioning of the brands is constantly reinforced in the consumer’s memory through the brands’ advertisements.

\(^4\) This conceptualization of a stable information set is also used by studies in experimental social psychology discussed in section 1. In these studies people’s “prior evaluations” about an object (i.e., their evaluations before they were given further information about the object in the experiment) are assumed to be stable throughout the length of the experiment.

\(^5\) We do not consider the impact of in-store activities like feature ads and displays on the quality beliefs of the consumer. This is because Mehta, Rajiv and Srinivasan (2003) have shown that presence of feature ads and displays have a negligible impact on consumer’s quality evaluations.
due to forgetting and decays completely after long periods of non-consumption of the brand. The implication of the last assumption is that if a consumer does not consume a brand for a long stretch of time, her evaluation of the quality of that brand would converge to her evaluation based on the positioning set.

(iv) On a purchase occasion, the consumer first constructs her quality evaluations of all brands corresponding to her consumption set. She does so by first recalling her quality evaluations of all brands based on her prior consumption set that she had constructed during her prior purchase occasion in the category. Following that she updates these recalled quality evaluations in a Bayesian fashion with the consumption signal she might have received if she purchased that brand on the prior purchase occasion. The Bayesian updating mechanism is discussed in section 2.21.

(v) While the consumer perfectly recalls the consumption signal of the brand she purchased on her prior purchase occasion in the category\(^6\), she imperfectly recalls her quality evaluations of all brands (based on the second information set) that she had constructed on her prior purchase occasion. The imperfect recall of prior quality evaluations based on prior consumption experiences is modeled by assuming that the consumer recalls her prior evaluations with an added noise. Further, the extent of the added noise in the prior quality evaluations is modeled by assuming that the extent of forgetting increases with time in a negative exponential fashion. The characterization of the imperfect recall of prior quality evaluations of all brands is discussed in section 2.21.

(vi) After having constructed her quality evaluations based on her consumption set, she updates them with her quality evaluations based on her positioning set to form her overall quality evaluations of all the brands. The consumer uses these overall quality evaluations to define her utility for each brand. On a given purchase occasion, the consumer observes the prices of all brands and chooses the brand that gives the highest expected utility.

2.2 Utility Specification and Evolution of Consumer’s Overall Quality Evaluations

Consider a product category with \(j = 1, \ldots, J\) brands. 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 overall quality evaluation, \(q_{ij,t}\), and price, \(p_{ij,t}\), as follows:

\[ u_{ij,t} = \alpha_j q_{ij,t} + \beta_j p_{ij,t} + \epsilon_{ij,t} \]

\(6\) We assume that the consumer does not forget the information that she learns about the brand’s quality from consumption of the brand she purchased on the last purchase occasion. This is especially true if we consider product categories like liquid detergents or ketchup where the consumer consumes the product throughout the purchase interval. For such product categories, it is reasonable to assume that the consumer would have consumed the last purchased brand fairly recently before making the next purchase in that product category.
\[ U_{ij,t} = \theta_i q_{ij,t} - p_{ij,t} \]  

(1)

The parameter \( \theta_i \) denotes consumer \( i \)'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’. \(^7\)

We assume consumer \( i \)'s true quality evaluation of given brand \( j \) to be \( q_{ij} \). We further assume that consumers are uncertain about their true quality evaluations of all brands. This is modeled by positing that consumer \( i \) on any purchase occasion \(^8\) \( t \), instead of being aware of the true quality of brand \( j \), holds only subjective overall quality evaluations that are given by

\[ q_{ij,t} \sim N(\omega_{ij,t},\sigma_{ij,t}^2), \quad \forall i, \forall j. \]

(2)

where \( \omega_{ij,t} \) denotes consumer \( i \)'s estimate of the expected overall quality of brand \( j \) at purchase occasion \( t \) and \( \sigma_{ij,t}^2 \) denotes the extent of consumer uncertainty about brand \( j \)'s quality at purchase occasion \( t \).

We assume that the consumer constructs her overall evaluations of the quality of any brand \( j \) at occasion \( t \) by mixing her quality evaluations from two mutually exclusive information sets: \( M_{ij,t} \) and \( C_{ij,t} \). The first information set, \( M_{ij,t} \), refers to consumer \( i \)'s positioning set for brand \( j \) at purchase occasion \( t \). As discussed before, we assume the positioning set, \( M_{ij,t} \), to be time invariant. We represent consumer \( i \)'s evaluation of brand \( j \)'s quality based on her information set, \( M_{ij,t} \), as \( q_{m,ij,t} \sim N(\omega_{m,ij,t},\sigma_{m,ij,t}^2) \). The consumption set, \( C_{ij,t} \), represents the cumulative information that consumer \( i \) has learned about the true quality of brand \( j \) through its consumption prior to purchase occasion \( t \). We represent consumer \( i \)'s evaluations about brand \( j \)'s quality at purchase occasion \( t \) based on the consumption set \( C_{ij,t} \) as \( q_{c,ij,t} \sim N(\omega_{c,ij,t},\sigma_{c,ij,t}^2) \). At the beginning of consumer \( i \)'s purchase history at \( t = 0 \), her consumption set is empty. Therefore, the consumer’s overall evaluation of brand \( j \)'s quality at purchase occasion \( t = 0 \) would be those based on her positioning set, \( M_{ij,0} \). On any other purchase occasion \( t \), the consumer would construct her overall quality evaluations of brand \( j \) from her total information set: \( \{M_{ij,t}, C_{ij,t}\} \). Since the two information sets are assumed to be mutually exclusive, the mean and variance of her overall quality evaluations at purchase occasion \( t \) would be given as

---

\(^7\)Please see R.D. Willig (1978) and W.M. Hanemann (1984) for additional details.

\(^8\) Note that the symbol ‘\( t \)’ does not represent the purchase occasion in terms of real time units. It should be interpreted as consumer \( i \)'s \( i^{th} \) purchase incidence in the category.
\[ \omega_{ij,t} = \left( \frac{\omega_{c,ij,t} + \omega_{y,0}}{\sigma_{\omega,ij,0}^2 + \frac{1}{\sigma_{\omega,ij,t}^2} + \frac{1}{\sigma_{\omega,0}^2}} \right)^{-1} \text{ and } \frac{1}{\sigma_{\omega,ij,t}^2} = \frac{1}{\sigma_{\omega,0}^2} + \frac{1}{\sigma_{\omega,ij,0}^2} \]  \tag{3}

On purchase occasion \( t \), consumer \( i \) updates her quality evaluations of brand \( j \) from her consumption set \( C_{ij,t} \). She does so by first recalling her quality evaluations of brand \( j \) that she had constructed from her consumption set \( C_{ij,t-1} \) on her previous purchase occasion \( t-1 \). Following that, she updates it with the information that she would have learned about brand \( j \)'s quality from its consumption if she purchased it on the prior purchase occasion. We assume that the consumption experience of brand \( j \) provides only a noisy signal of consumer \( i \)'s true quality evaluation of brand \( j \). Let \( \hat{\lambda}_{ij,t-1} \) be the noisy quality signal received by consumer \( i \) from brand \( j \)'s purchase at occasion \( t-1 \) with

\[ \hat{\lambda}_{ij,t-1} = q_{ij} + \eta_{ij,t-1} \]  \tag{4}

where \( \eta_{ij,t-1} \) denotes the noise associated with consumption experience, and is independent and identically distributed across all consumers, all brands and all time periods as \( \eta_{ij,t-1} \sim N(0, \sigma_{\eta}^2) \). \footnote{Note that } \sigma_{\eta}^2 \text{ is a measure of the informativeness of consumption experience: } \sigma_{\eta}^2 = 0 \text{ corresponds to the product category being classic experience goods so that the consumer gets to learn the “true” quality after a single consumption experience.}

As discussed in section 2.1, we assume that consumer \( i \) receives the consumption signal \( \hat{\lambda}_{ij,t-1} \) just before making her next shopping trip at occasion \( t \). This implies that she would perfectly recall the consumption signal \( \hat{\lambda}_{ij,t-1} \) while constructing her new evaluations of brand \( j \) at purchase occasion \( t \). On the other hand, since time could have lapsed since she last constructed her quality evaluations of brand \( j \) at occasion \( t-1 \), she would imperfectly recall her prior quality evaluations, \( q_{c,ij,t-1} \), at occasion \( t \). We represent her imperfectly recalled quality evaluation of brand \( j \) as \( q_{Re,ij,t-1} \) that is distributed as \( N(\hat{\lambda}_{ij,t-1}, \sigma_{Re,ij,t-1}^2) \). Using the fact that normal density is self conjugate, the mean and variance of consumer’s quality evaluations of brand \( j \) at purchase occasion \( t \), \( q_{c,ij,t} \), will be related to the mean and variance of her imperfectly recalled evaluations, \( q_{Re,ij,t-1} \) as

\[ \omega_{c,ij,t} = \frac{\sigma_{Re,ij,t-1}^2 + d_{ij,t-1} \hat{\lambda}_{ij,t-1}}{\sigma_{\eta}^2} \text{ and } \frac{1}{\sigma_{\omega,ij,t}^2} = \frac{1}{\sigma_{Re,ij,t-1}^2} + \frac{1}{\sigma_{\eta}^2} \]  \tag{5}

where \( d_{ij,t-1} \) is an indicator variable that takes the value of 1 if brand \( j \) was purchased by consumer \( i \) at purchase occasion \( t-1 \). In order to fully specify the evolution of consumer’s quality...
evaluations over time, we will characterize the mean and variance of her recalled evaluations, $q_{rc,ij,t-1}$, as a function of the mean and variance of her prior evaluations of the prior evaluations, $q_{c,ij,t-1}$, in the following section.

2.2.1 Characterization of Consumer’s Recalled Quality Evaluations

In order to model the imperfect retrieval of consumer’s prior evaluations of a brand based on her consumption set, we will define the consumer’s quality evaluations at occasion $t-1$, $q_{c,ij,t-1}$, in terms of just one hypothetical aggregate consumption signal, $S_{g,ij,t-1}$, that consumer $i$ would have received just before her purchase occasion $t-1$ on the basis of which she formed her evaluations of brand $j$ as $q_{c,ij,t-1}$. The hypothetical aggregate consumption signal would be $S_{g,ij,t-1} = \omega_{c,ij,t-1}$ that is distributed around her true quality evaluation of brand $j$ as $N(q_j, \sigma_{c,ij,t-1}^2)$ . Note that the hypothetical aggregate signal, $S_{g,ij,t-1}$, acts as a proxy for consumer $i$’s quality evaluation of brand $j$ based on her consumption set occasion $t-1$. Therefore, instead of modeling the imperfect recall of $q_{c,ij,t-1}$, we will model the imperfect recall of the aggregate signal, $S_{g,ij,t-1}$ at occasion $t$ in order to derive the specification of the mean and variance of the recalled quality evaluations.

If consumer $i$ were to perfectly recall the hypothetical aggregate consumption signal $S_{g,ij,t-1}$ at purchase occasion $t$, her recalled quality evaluations, $q_{rc,ij,t-1}$, would be identical to her prior quality evaluations, $q_{c,ij,t-1}$ . On the other hand, if consumer $i$ were to imperfectly recall the aggregate signal $S_{g,ij,t-1}$ at purchase occasion $t$, we posit that she would recall it with an added noise, $\varphi_{g,ij,t-1}\hat{v}_{g,ij,t-1}$, such that the recalled prior aggregate consumption signal, $S_{r,ij,t-1}$, at occasion $t$ would be

$$S_{r,ij,t-1} = S_{g,ij,t-1} + \varphi_{g,ij,t-1}\hat{v}_{g,ij,t-1}$$

In equation (6), $\hat{v}_{g,ij,t-1}$ is a realized value of a standard normal random variable that is independent across all consumers, all brands and all purchase occasions. Further, $\varphi_{g,ij,t-1}$ represents the extent of noise that is added due to the imperfect retrieval of the aggregate signal, $S_{g,ij,t-1}$.

As shown in the Technical Supplement, the mean and variance of consumer $i$’s recalled prior quality evaluations of brand $j$ at purchase occasion $t$ based on the aggregate recalled signal $S_{r,ij,t-1}$ would be

$$\omega_{rc,ij,t-1} = \omega_{c,ij,t-1} + \varphi_{g,ij,t-1}V_{g,t-1}$$

$$\sigma_{rc,ij,t-1}^2 = \sigma_{c,ij,t-1}^2 + \phi_{g,ij,t-1}^2$$

(7)
Equation (7) gives the specification of consumer’s recalled quality evaluation based on her consumption set at purchase occasion \( t \) as a function of her quality evaluation based on her consumption set at purchase occasion \( t-1 \) and her extent of forgetting, \( \varphi_{ij,t-1} \). Next, we will specify the extent of forgetting, \( \varphi_{ij,t-1} \), in terms of consumer i’s characteristics and the real time lapsed between occasions \( t-1 \) and \( t \). In order to do so, we will use the result of the experimental findings in forgetting that suggest the extent of forgetting of prior evaluations to be an increasing and concave function of the inter-purchase time.

Recall that the extent of forgetting has been conceptualized in the forgetting literature (Rubin and Wenzel 1996) as the extent of decay in the precision of the prior evaluations. Further, the literature suggests that the extent of decay is captured reasonably well by an exponential decay function of time. Note that in our context, the precision of the prior evaluation is simply the inverse of the variance of the prior evaluations, \( \sigma_{c,ij,t-1}^2 \). Using an exponential decay function, we get the following specification of the extent of forgetting noise, \( \varphi_{ij,t-1} \), as a function of the time lapsed between purchase occasions \( t-1 \) and \( t \)

\[
\varphi_{ij,t-1}^2 = \sigma_{c,ij,t-1}^2 \left( e^{b_{ij}^w w_{ij,t-1,t}} - 1 \right).
\]  

(8)

Therefore, the mean and variance of the recalled quality evaluation, \( q_{R,c,ijb} \), would be related to the mean and variance of the prior evaluation, \( q_{c,ijb} \), as

\[
\omega_{R,c,ijb} = \omega_{c,ijb} + \hat{y}_{ij,t-1} \sigma_{c,ij,t-1} \sqrt{e^{b_{ij}^w w_{ij,t-1,t}} - 1}
\]

\[
\sigma_{R,c,ijb}^2 = \sigma_{c,ijb}^2 e^{b_{ij}^w w_{ij,t-1,t}}
\]  

(9)

In equation (9), \( w_{i,t-1,t} \) represents the time lapsed in weeks between the purchase occasions \( t-1 \) & \( t \), and \( b_i \) represents consumer i’s extent of forgetting of her prior evaluation per week for any of the brands in the product category. Note that the consumer’s extent of forgetting depends on the real time lapsed between purchase occasions \( t-1 \) and \( t \). Since purchase intervals need not be the same during the consumer’s purchase history, the consumer’s extent of forgetting can vary across between purchase occasions over her purchase history. For simplicity, we assume consumer i’s extent of forgetting per week, \( b_i \), to be the same across all the brands in the product category.

Equation (9) generates the following implications regarding the impact of the parameter \( b_i \) on the consumer’s extent of forgetting: (i) as \( b_i \) increases, the extent of uncertainty in the recalled evaluations increases. In particular, \( b_i = 0 \) implies that the consumer would perfectly recall her prior quality evaluations and \( b_i = \infty \) implies that the consumer would completely forget her prior...
quality evaluations.\(^{10}\) (ii) Since we have assumed an exponential decay function to capture the loss in precision of recalled evaluations, it implies that it would take \(\frac{\ln(2)}{b_i}\) weeks for the precision in the recalled evaluations to decay to half of their prior value.

Note that in our formulation, forgetting impacts both the mean and the variance of the recalled quality evaluation. Further, the prior mean quality evaluation gets distorted as a result of forgetting because of the inclusion of the random error term, \(\hat{\nu}_{tij\lt-1}\), whose magnitude depends on the extent of forgetting that occurs between purchase occasion \(t\)-1 and \(t\). Therefore, unlike the conceptualization of forgetting in Keon (1980), in our formulation forgetting impacts both the mean and the precision of recalled quality evaluations. Further, unlike the conceptualization of forgetting in Mullainathan (2002), in our formulation forgetting causes consumers to recall distorted versions of their prior evaluations.

In summary, we have specified the consumer’s recalled quality evaluations based on her consumption set as a function of her prior evaluations based on her prior consumption set and her prior inter-purchase time in this section. In the next section, we will specify the evolution of consumer’s overall quality evaluations by combining her evaluations from her consumption and positioning sets.

### 2.22 Evolution of Overall Quality Evaluations from Consumer’s Perspective

The specification of the mean and variance of consumer’s overall quality evaluations, \(q_{ij\lt}\), can be derived by substituting the expressions for the mean and variance of the recalled evaluations, \(q_{rc\lt}\), in equation (9) into equations (3) and (5) as follows

\[
\alpha_{ij\lt} = \frac{f\alpha_{\lt-1} + (1-f)\alpha_{ij\lt-1} + (1-f)\alpha_{ij\lt-1}}{f\alpha_{\lt-1} + (1-f)\alpha_{ij\lt-1} + (1-f)\alpha_{ij\lt-1}} + \frac{d_{ij\lt-1}^2}{f\alpha_{\lt-1} + (1-f)\alpha_{ij\lt-1} + (1-f)\alpha_{ij\lt-1}} + \frac{\nu_{\lt-1}^2}{f\alpha_{\lt-1} + (1-f)\alpha_{ij\lt-1} + (1-f)\alpha_{ij\lt-1}}
\]

\(\alpha_{ij\lt} = f\alpha_{\lt-1} + (1-f)\alpha_{ij\lt-1} + d_{ij\lt-1}\)

where \(f = \exp(-b_i\nu_{\lt-1})\), \(\alpha_{ij\lt-1} = \sigma_q^2 / \sigma_{\lt-1}^2\) is the precision of consumer \(i\)’s evaluation of brand \(j\) at time \(t\), \(\alpha_{ij\lt-1} = \sigma_q^2 / \sigma_{\lt-1}^2\) is the precision of consumer’s evaluations at the beginning of her

---

\(^{10}\) This ensures that as the time lapsed between purchase occasions \(t\)-1 and \(t\) becomes infinite, the recalled information set at purchase occasion \(t\) corresponding to past consumption experiences till occasion \(t\)-1 would decay completely and the consumer’s recalled prior evaluations based on past consumption experiences till purchase occasion \(t\)-1 would become non-informative.
consumption history at time $t=0$. We can interpret $\alpha_{ij,0}$ as the consumer’s information about the quality of brand $j$ in her positioning set and $\alpha_{ij,t}$ as the consumer’s total information about brand $j$’s quality in her positioning and consumption sets at occasion $t$.

It is important to note that equation (10) represents the evolution of quality evaluations from consumer $i$’s perspective. Note that consumer $i$ deterministically knows the value of $\omega_{ij,t}$ since she observes the consumption signal $\hat{t}_{ij,t-1}$ and knows the value of the mean of the recalled prior evaluation, $\omega_{Re,ij,t-1}$ (which is the sum of $\omega_{c,ij,t-1}$, and the realization of the random noise, $\hat{t}_{ij,t-1}\sigma_{c,ij,t-1}(\sqrt{e^{b_{ij,t-1}}}-1)$). Having specified the evolution of consumer’s overall quality evaluations in equation (10), we will next discuss the insights that it generates regarding the impact of forgetting on the decay of quality evaluations over time. The analytical insights are summarized in the following two propositions:

**Proposition 1:** As long as $b_i > 0$, $\sigma_\eta^2 > 0$, and $w_{i,t-1,t} > 0 \forall t$, consumer $i$ can never learn about her true quality evaluation of a brand even after repeat purchasing it infinite times.

Proposition 1 states that if consumer’s extent of forgetting of her recalled evaluations is greater than zero (as captured by $b_i w_{i,t-1,t} > 0$) and if her consumption experiences do not fully reveal her true quality evaluations of the brand (as captured by $\sigma_\eta^2 > 0$), she can never realize her true quality evaluation even after consuming the brand infinite times. This is because the consumer’s recalled evaluation of a brand on any purchase occasion will have an added noise that would not diminish to zero even after infinite consumption experiences - as a result, the uncertainty in her quality evaluations will always persist. The asymptotic level of uncertainty in consumer’s quality evaluation of the brand after repeated consumption depends on her extent of forgetting, $b_i$, and the distribution of her inter-purchase times. If the extent of consumer’s forgetting, $b_i$, is small or her purchase intervals are short, then the asymptotic level of uncertainty in her quality evaluation will be small too. Further, if either $b_i = 0$ or $w_{i,t-1,t} = 0 \forall t$, the consumer would realize her true quality evaluation, $q_{ij}$, after infinite consumption experiences.

The next proposition summarizes the factors that affect the extent of decay of consumer’s mean quality evaluations of a brand in absence of its consumption.

---

11 The proofs of the propositions are given in the Appendix
Proposition 2: In absence of purchase of a brand at purchase occasion $t$ and thereafter, the consumer’s mean quality evaluation of that brand reverses back with time to her mean quality evaluation at the beginning of her purchase history. Further,

(i) The mean quality evaluations decay in a sigmoid fashion if the information in the consumer’s consumption set, $\alpha_{i,t} - \alpha_0$, is greater than the information in her positioning set, $\alpha_0$.

(ii) The mean quality evaluations decay in a convex fashion if the information in the consumer’s consumption set, $\alpha_{i,t} - \alpha_0$, is smaller than the information in her positioning set, $\alpha_0$.

(iii) The extent of reversion of the mean quality evaluation over time is an increasing function of consumer’s extent of forgetting per week, $b_i$;

(iv) The extent of reversion of the mean quality evaluation is a decreasing function of the ratio of the information in the consumer’s consumption set $\alpha_{i,t} - \alpha_0$ to the information in her positioning set, $\alpha_0$.

The implications of proposition 2 are shown in Figures 1A and 1B. Figure 1A shows the decay of consumer’s mean quality evaluations of a brand in time periods following purchase occasion $t$ for a range of values of the extent of forgetting $b_i$. Note that the extent of decay of overall quality evaluations increases with the extent of forgetting $b_i$. Figure 1B shows the decay of mean quality evaluations for a range of values of the ratio of the information in consumer’s consumption set to the information in her positioning set, $k_{i,t} = \frac{\alpha_{i,t} - \alpha_0}{\alpha_0}$. Note that the extent of decay of mean overall quality evaluations decreases with $k_{i,t}$. This implies that the greater the extent of the prior learning done through consumption (relative to the prior learning done at the beginning of the consumption history), the smaller is the decay of mean quality evaluations with time. This result conforms to the experimental findings of Wilson and Miller (1968) and Johnson and Watkins (1971) who report that the extent of decay of people’s overall evaluations varies inversely with their extent of prior learning about the attitude object.

Further observe in Figure 1B that the shape of the decay of overall quality evaluations is convex for values of $k_{i} < 1$ and is sigmoid for values of $k_{i} > 1$. Comparing this result with the results of the experiments discussed in section 1, we see that (i) the shape of decay of mean quality evaluations for $k_{i} > 1$ is similar to the those reported in the experimental findings of Papageorgis (1963) and Watts and McGuire (1964); (ii) the shape of decay of mean quality evaluations for $k_{i} < 1$ is similar to those reported in the experimental findings of Collins and Hoyt.
(1972) and Ronis et. al. (1977). Therefore, our model is able to explain the two different sets of results regarding the shape of the decay of evaluations over time as reported in previous studies in experimental social psychology.

2.3 Econometric Specification of Consumer’s Quality Evaluations

As discussed in section 2.2.1, the consumer observes the realization of the random signals \( \{ \hat{\lambda}_{ij,s} \}_{s=1}^{t-1} \) and hence, she knows deterministically the values of the mean of her overall quality evaluation, \( \omega_{ij,t} \), for all brands \( j \) at any purchase occasion \( t \). But the analyst does not observe the realizations of the random quality draws \( \{ \hat{\lambda}_{ij,s} \}_{s=1}^{t-1} \). Further, the analyst does not observe the realizations of the random noise terms, \( \{ \hat{\nu}_{ij,s} \}_{s=1}^{t-1} \), that get added due to the imperfect recall of prior evaluations. Therefore, from the analyst’s perspective, both \( \{ \hat{\lambda}_{ij,s} \}_{s=1}^{t-1} \) and \( \{ \hat{\nu}_{ij,s} \}_{s=1}^{t-1} \) are random variables and hence, the analyst’s estimate of \( \omega_{ij,t} \) (which we denote by \( \tilde{\omega}_{ij,t} \)) is a random variable too.

From the analyst’s perspective, the random errors, \( \{ \hat{\nu}_{ij,s} \}_{s=1}^{t-1} \), will be independently distributed as \( \mathcal{N}(0,1) \). Further, from the analyst’s perspective at purchase occasion \( t \), \( \lambda_{ij,t-1} \) will be a random variable as given in equation (4) by \( \lambda_{ij,t-1} \sim \mathcal{N} \left( q_{ij}, \sigma_{ij}^2 \right) \). But note that at occasion \( t \), neither the consumer nor the analyst knows the true quality evaluation, \( q_{ij} \), of brand \( j \). Hence, the analyst can not assume \( q_{ij} \) to be the mean of the quality signal, \( \lambda_{ij,t-1} \). Instead, the analyst will use his most accurate belief (as per his information set prior to purchase occasion \( t \)) about the true quality evaluation \( q_{ij} \) to be the mean of the signal \( \lambda_{ij,t-1} \). As shown in the Technical Supplement, the analyst’s information set about brand \( j \)'s true quality prior to consumer \( i \)'s purchase occasion \( t \) is \( \{ \hat{\omega}_{ij,s}, \sigma_{ij,s}^2 \}_{s=1}^{t-1} \). Hence his most accurate belief of the true quality of brand \( j \) would be \( q_{ij,r} \sim \mathcal{N}(\hat{\omega}_{ij,r}, \sigma_{ij,r}^2) \) such that \( r = \arg\min_s \left\{ \sigma_{ij,s}^2 \right\}_{s=1}^{t-1} \). Therefore, it follows that from the analyst’s

\(^{12}\)Note that this approach for characterizing the signal, \( \lambda_{ij,t-1} \), from the analyst’s perspective is consistent with Jovanovic (1979), Miller (1984) and Mehta, Rajiv and Srinivasan (2003).
perspective, \( \lambda_{ij,t-1} \) will be distributed as
\[
\lambda_{ij,t-1} = \tilde{\omega}_{ij,t} + \xi \sqrt{\sigma^2 + \sigma^2_{\omega,t}}
\]
where \( \xi \) is a standard normal random variable\(^{13}\).

Substituting the analyst’s estimates for \( \tilde{\lambda}_{ij,t-1} \) and \( \tilde{\nu}_{ij,t-1} \) into equation (10), we get the analyst’s estimate of the mean and precision of consumer \( i \)'s overall quality evaluation of brand \( j \) at occasion \( t \) as
\[
\begin{align*}
\bar{\omega}_{ij,t} & = \frac{f \alpha_{ij,t-1} \tilde{\omega}_{ij,t-1}}{f \alpha_{ij,t-1} + (1 - f) \alpha_{ij,0} + d_{ij,t-1}} + \frac{(1 - f) \tilde{\omega}_{ij,0} \alpha_{ij,0}}{f \alpha_{ij,t-1} + (1 - f) \alpha_{ij,0} + d_{ij,t-1}} + \\
& + \frac{d_{ij,t-1} \left( \tilde{\omega}_{ij,t} + \xi \sqrt{1 + \alpha_{ij,t}} \right)}{f \alpha_{ij,t-1} + (1 - f) \alpha_{ij,0} + d_{ij,t-1}} + \frac{\nu_{ij,t-1} \sqrt{(\alpha_{ij,t-1} - \alpha_{ij,0}) f (1 - f)}}{f \alpha_{ij,t-1} + (1 - f) \alpha_{ij,0} + d_{ij,t-1}} \\
\alpha_{ij,t} & = f \alpha_{ij,t-1} + (1 - f) \alpha_{ij,0} + d_{ij,t-1}
\end{align*}
\] (11)

In equation (11), the analyst’s estimate of the prior mean quality belief, \( \tilde{\omega}_{ij,t-1} \) is a truncated normal random variable that is subject to restrictions imposed in the error space due to consumer \( i \)'s brand choices on purchase occasions \( s=1 \) till \( t-1 \)^{14}. Further, \( \xi \) and \( \nu_{ij,t-1} \) are standard normal random variables that represent the analyst’s uncertainty in predicting the extent of consumer’s learning from the prior consumption and her extent of forgetting of prior evaluations respectively. Thus our model allows for a very flexible learning and forgetting mechanism where both can have favorable or an unfavorable impact on consumer’s quality evaluations. This specification is different from reduced form models that use the exponentially smoothed brand loyalty (e.g., Guadagni and Little 1983). Even though the brand loyalty term does implicitly capture learning from consumption and forgetting of prior evaluations (since the brand loyalty increases with repeated consumption and decreases in absence of consumption), it only allows for a unidirectional and a deterministic change in the consumer’s evaluations. Further the brand loyalty only captures the consumer’s state dependence and does not generate a correlation in random terms in brand evaluations across purchase occasions. On the other hand our conceptualization not only generates a last period’s state dependence as a result of learning from prior consumption, but also generates a correlation in random errors. We will further discuss these implications in section 2.33.

---

\(^{13}\) Since we can only estimate \( \alpha_{ij,t-1} = \sigma^2_{\omega,t-1}/\sigma^2_{\omega,t-1} \), for identification purposes, we set the variance of consumption signal to be one i.e. \( \sigma^2_{\omega,t-1} = 1 \).

\(^{14}\) These restrictions are discussed in detail in the Technical Supplement.
2.31 Stochastic Specification of Consumer’s Brand Choice Decision

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. Further, we assume that the value of the precision of the subjective quality beliefs of all the brands at the beginning of the consumer’s purchase history is $\alpha_0$. In other words, $\tilde{\omega}_{ij,0} = 0$ and $\alpha_{ij,0} = \alpha_0$ for all consumers $i$ and brands $j$. It follows that consumer $i$’s expected indirect utility for brand $j$ at purchase occasion $t$ would be

$$EU_{ij,t} = \frac{\theta_i \alpha_{ij,t-1} \tilde{\omega}_{ij,t-1} + \theta_i d_{ij,t-1} \lambda_{ij,t-1}}{f_{ij,t-1} + (1-f)\alpha_0 + d_{ij,t-1}} + \frac{\theta_i v_{ij,t-1} \sqrt{[\alpha_{ij,t-1} - \alpha_0]} f(1-f)}{f_{ij,t-1} + (1-f)\alpha_0 + d_{ij,t-1}} - p_{ij,t}$$

(12)

Since $\tilde{\omega}_{ij,t}$’s are truncated normal random variables, the probability that consumer $i$ selects brand $j$ at purchase occasion $t$ is given by

$$\Pr(d_{ij,t} = 1) = \Pr(\theta \tilde{\omega}_{ij,t} - p_{ij,t} > \theta \tilde{\omega}_{ikt} - p_{ikt}, \forall k \neq j)$$

$$= \int_{-\infty}^{p_{ij,t}} \cdots \int_{-\infty}^{p_{ij,t}} \text{Truncated} \phi \left(\tilde{\omega}_{ij,t}^{1}, \cdots, \tilde{\omega}_{ij,t}^{1}, \tilde{\omega}_{ij,t}^{1}, \cdots, \tilde{\omega}_{ij,t}^{N} \right) \prod_{h=1}^{N} d\tilde{\omega}_{ij,t}^{h}$$

(13)

where the variables $\tilde{\omega}_{ij,t}^{h}$ and $p_{ij,t}^{h}$ as defined as: $\tilde{\omega}_{ij,t}^{h} = \tilde{\omega}_{ih,t} - \tilde{\omega}_{ij,t}$ and $p_{ij,t}^{h} = p_{ih,t} - p_{ij,t}$.

2.32 Controlling for Unobserved Heterogeneity

In the proposed formulation, since consumers go through different idiosyncratic consumption experiences and different extents of forgetting with time, the quality evaluations are not the same across all consumers. This allows for a richer specification of the intercept heterogeneity as compared to the extant literature (Chintagunta, Vilcassim and Jain 1991). To account for unobserved heterogeneity in consumer’s marginal willingness to pay for quality, we assume $\theta_i$ to be gamma distributed across the consumer population with mean $\bar{\theta}$ and variance $\sigma_{\theta}^2$. Further, to account for unobserved heterogeneity in the consumer’s extent of forgetting per week, we assume that $b_i$ is gamma distributed across the consumer population with mean $\bar{b}$ and

---

15 Since our data presents a left truncation problem, we can not assume the mean and precision of the consumer’s quality evaluations on the first purchase occasion to be $q_{ij,0} \sim N(\omega_0, \alpha_0^{-1})$. In our empirical analysis, we use an initialization sample to estimate consumer’s initial quality evaluations and use them for the estimation sample.
2.33 Implications on Consumer’s Habit Persistence and State Dependence

Habit persistence has been defined as the impact of consumer’s prior propensity to purchase a brand on her current purchase probability of that brand (Heckman 1981; Roy, Chintagunta and Haldar 1996). This is traditionally captured by the serial correlation in consumer’s utility of a brand on two consecutive purchase occasions. Note that in our formulation, the coefficient of $\widehat{\omega}_{j,t-1}$ in the consumer’s utility for brand $j$ at occasion $t$, that is

\[
\frac{\theta e^{b_{i,t-1}} \alpha_{j,t-1}}{e^{b_{i,t-1}}(\alpha_{j,t-1} - \alpha_{j,0}) + \alpha_{j,0} + d_{j,t-1}},
\]

captures the correlation in the stochastic terms in the consumer i’s utility of brand $j$ at purchase occasions $t-1$ and $t$. Therefore, we can interpret the coefficient of $\widehat{\omega}_{j,t-1}$ as the habit persistence in consumer i’s utility of brand $j$ at purchase occasion $t$.

Consumer’s state dependence has been defined as the impact of the past purchase of a brand (through its consumption experience) on the current choice probability of that brand (Heckman 1981; Roy, Chintagunta and Haldar 1996). In our model, state dependence can be interpreted as the added information that the consumer gets about the quality of a brand from its consumption in the prior time period (Erdem and Keane 1996). Therefore, the coefficient of $d_{j,t-1}$ in equation (12), that is

\[
\frac{\theta e^{b_{i,t-1}} \lambda_{j,t-1}}{e^{b_{i,t-1}}(\alpha_{j,t-1} - \alpha_{j,0}) + \alpha_{j,0} + d_{j,t-1}},
\]

can be interpreted as the state dependence in consumer i’s utility of brand $j$ at occasion $t$.

The above discussion implies that forgetting impacts habit persistence and last period’s state dependence in the consumer’s utility as follows: (i) Forgetting decreases the impact of the prior mean quality evaluation of brand $j$, $\widehat{\omega}_{j,t-1}$. It increases the impact of the mean quality evaluation of brand $j$ formed prior to any consumption experience, $\widehat{\omega}_{j,0}$. This implies that forgetting leads to a decrease in habit persistence in the consumer’s utility. (ii) Forgetting increases the level of uncertainty in the prior recalled evaluation of brand $j$. This leads to an increase in the impact of learning from consumption of the brand $j$ purchased on the prior purchase occasion. This implies that forgetting leads to an increase in last period’s state dependence in the consumer’s utility.

The expressions for state dependence and habit persistence also imply that: (i) the heterogeneity in both state dependence and habit persistence results from the heterogeneity in variance $\sigma^2_h$. This ensures that the marginal willingness to pay for quality and the extent of imperfect recall are positive.
quality sensitivity, $\theta_i$, and the extent of forgetting, $b_i$, across the consumer population. (ii) The inter-temporal variation in state dependence and habit persistence for a given consumer results from the variation in her inter-purchase time, $w_{i,t-1,t}$, and the variation in the precision of her quality evaluations, $\alpha_{i,j,t-1}$, through her purchase history. We further expand on the second implication in the following two propositions\(^\text{16}\).

**Proposition 3:** The habit persistence in consumer’s purchase behavior at purchase occasion $t$ decreases with the length of the inter-purchase time between occasions $t-1$ and $t$ as a result of forgetting. Further, the extent of decrease in habit persistence with time is a decreasing function of the cumulative information that the consumer had learned from her consumption experiences prior to occasion $t-1$.

The intuition behind proposition 3 is as follows: as the length of the prior purchase interval increases, the extent of forgetting of the prior evaluations would increase. It follows that the contribution of the consumer’s prior evaluations to her present utility would decrease. This implies that the habit persistence in the consumer’s purchase behavior would decrease. Further, as discussed in proposition 2, the greater the information in the consumer’s consumption set, the smaller will be the extent of decay of the prior quality evaluations. As a result, the larger will be the habit persistence in her purchase behavior.

**Proposition 4:** Consumer’s state dependence in her purchase behavior at purchase occasion $t$ increases with the length of her inter-purchase time between occasions $t-1$ and $t$ as a result of forgetting. Further, the extent of increase of the state dependence with the length of the inter-purchase time is a decreasing function of the cumulative information that the consumer had learned from her consumption experiences of that brand prior to occasion $t-1$.

The intuition behind proposition 4 is as follows: as the length of the inter-purchase time increases, the consumer’s uncertainty in her recalled evaluations also increases. Therefore, the consumer’s extent of learning from consuming a brand that she purchased on the prior occasion would increase - which implies that the state dependence in the consumer’s utility would also increase. Further, the smaller the information learned by the consumer from consumption experiences prior to occasion $t-1$, the smaller will be the extent of forgetting of the recalled evaluations and hence, the greater will be the impact of last period’s state dependence in her present utility. Note that this result is contrary to the empirical findings of Seetharaman, Ainslie and Chintagunta (1999) who observe that state dependence is a decreasing function of the prior

\(^{16}\) The proofs of the propositions are given in the Appendix
inter-purchase time. The difference in the two results can be explained as follows. Unlike the proposed model, the model proposed by Seetharaman, Ainslie and Chintagunta (1999) does not explicitly consider the change in consumer’s habit persistence over time. Therefore, it is possible that the model proposed by Seetharaman, Ainslie and Chintagunta (1999) might give biased estimates for the change in state dependence with time.

3. DATA, ANALYSIS AND DISCUSSION

3.1 Data

For model calibration and analysis, we use the ERIM data set for liquid detergents. We have taken 4 national brands for the analysis: Wisk, Tide, Era and Surf. These four brands account for a total of 81% of the market share in this product category. The data set comprises a random sample of 400 households with purchase observations extending from the 25th week of 1986 to the 34th week of 1988. The minimum number of purchase 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 3592 purchase observations and the holdout sample had a total of 3019 purchase observations. In both the estimation and the holdout samples, we take the first five observations for each household to initialize consumers’ prior beliefs about the qualities of these brands and the store familiarity variable for the consumers. The summary statistics for the entire sample of 400 households are given in Table 1 and Exhibit 1.

Since our data presents a left truncation problem, we cannot assume the mean and precision of the consumer’s quality evaluations on the first purchase occasion to be 
\[ q_{ij,0} \sim N(\omega_0, \alpha_0^{-1}) \]. Therefore, we make the following corrections regarding the prior mean and precision of the quality evaluations for all brands and for all consumers: first, we assume that the initial precision of the quality evaluations for all brands and for all consumers in the estimation sample is \( \alpha_0 \) (and not \( \alpha_0 \)) and separately estimate the two. Second, we use a pre-estimation sample in which we assume that consumers have the same quality beliefs for all the brands at the beginning of their purchase histories, that is, \( q_{ij,0} \sim N(\omega_0, \alpha_0^{-1}) \forall i, j \). We calibrate the model for

\(^{17}\)But note that in our estimation results using the model proposed by Seetharaman, Ainslie and Chintagunta (1999) on liquid detergents, we find that the state dependence actually increases with the length of the inter-purchase time and the parameter capturing the change in state dependence with time is statistically significant.
the pre-estimation sample and calculate the posterior mean quality beliefs for all the brands and for every consumer at the end of their fifth purchase occasion. We then take these posterior mean quality beliefs as the initial prior mean beliefs for the estimation sample\(^{18}\). The mean posterior quality beliefs across all consumers for the pre-estimation sample for liquid detergents are given in Table 2.

### 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 two other competing models. We coded the program in MATLAB. We estimated the following parameters for the proposed model (referred to as Model I henceforth):

(i) Mean quality sensitivity parameter across the consumer population, \(\bar{\theta}\) (We assume quality sensitivity parameter \(\theta\) to be gamma distributed with mean \(\bar{\theta}\) and variance \(\sigma^2\theta\));

(ii) Variance of quality sensitivity across the population, \(\sigma^2\theta\).

(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, \(\alpha_s\);

(iv) Ratio of the noise in consumption signal to the information that could be gained at the beginning of her consumption history, \(\alpha_0\);

(v) Log of the mean value of the extent of decay per week of the information learned from prior consumption experiences across the consumer population, \(\bar{b}\) (We assume \(b_i\) to be gamma distributed with mean \(\bar{b}\) and variance \(\sigma^2_b\));

(vi) Log of the variance of the extent of decay per week of the information learned from prior consumption experiences across the consumer population, \(\sigma^2_b\).

The parameter estimates for the liquid detergent data set are given in Table 3 (Column 2). The parameters are by and large statistically significant. For this data set, we compare the predictive power and goodness-of-fit of the proposed specification (Model I) to two other competing specifications:

1) A variant of the proposed model (referred to as Model II henceforth) obtained by assuming that consumers perfectly recall their prior quality evaluations of all brands. The implied

---

\(^{18}\) Note that this procedure is similar to the one used by Mehta, Rajiv and Srinivasan (2003)
parametric restrictions are \( \bar{b} = 0 \) and \( \sigma_b^2 = 0 \). This model, however, allows for quality learning through consumption. The parameters that we estimate in Model II are as follows:

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

b) Variance of quality sensitivity across the population, \( \sigma^2 \);

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, \( \alpha \);

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

2) The model of state dependence proposed by Seetharaman, Ainslie and Chintagunta (1999) (referred to as SAC (1999) henceforth)\(^{19}\). This model is chosen for comparison because it accounts for inter-temporal and cross sectional variations in consumer’s state dependence. In order to ensure that the number of explanatory variables for SAC (1999) is same as that in Model I, we do not take the impact of feature ads and displays on the consumer’s purchase decision. The consumer’s utility for a brand is a function of its brand dummy, log of its price, the presence of purchase of that brand in the prior purchase occasion and the effect of the time lapsed since the last purchase on state dependence. The consumer heterogeneity is captured by assuming that all parameters are distributed normally across the consumer population. The parameter estimates for SAC (1999) for the liquid detergent data set are summarized in Table 4.

Table 5A summarizes the results from comparison of the proposed model (Model I) against the two competing models (Model II, SAC (1999)) for liquid detergents. We use the statistical test proposed by Singleton (1985) that is appropriate for comparing non-nested specifications in the Generalized Method of Moments (GMM) framework\(^{20}\). We report the test results for both estimation and holdout samples. We find that the test rejects Model II in favor of Model I for both estimation sample (p-value<0.05) and the holdout sample (p-value<0.02). This supports the inference that consumers do not perfectly recall their prior quality evaluations while making their brand choice decision. Similarly, the test rejects the SAC (1999) 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 SAC (1999), the Model I is a superior representation of consumer brand choice behavior for liquid

---

\(^{19}\) Note that the model proposed by Seetharaman, Ainslie and Chintagunta (1999) was simultaneously estimated on five product categories. In our analysis, we estimate their model on only one product category.

\(^{20}\) In the Singleton test (1985), the null hypothesis is that the competing model is the true model of consumer behavior. The test entails whether Model I significantly outperforms the competing model. The test statistic is distributed as \( \chi^2 \) with 1 d.o.f.
detergents. Table 5B 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.

3.3 Results for Liquid Detergents and Discussion of Insights

We summarize our discussion under five broad substantive issues:

**Perceived Quality Evaluations of the Brands:** From our parameter estimates in Table 2 we find that the mean perceived quality evaluations across the consumer population of Wisk, Tide, Era and Surf are \(^{21}\) 0.0110, 0.0203, 0.0103 and 0.0091 respectively for Model I. On the other hand, the mean perceived quality evaluations of the respective brands as predicted by Model II are 0.0431, 0.0624, 0.0469 and 0.0307. This implies that: (i) at the aggregate level, Model I predicts a smaller variance in the perceived quality evaluations across the brands as compared to Model II. (ii) Since the deviation of these average quality evaluations from the prior value of zero represents the aggregate learning that has been done from consumption experiences, it implies that the predicted extent of aggregate leaning done through consumption experiences is smaller for Model I as compared to Model II.

We further test the two implications at the household level for both Models I and II. In order to test the first implication, we first calculate the standard deviation in the perceived quality beliefs across the four brands for each purchase observation in the data. Following that, we average them across all purchase observations for both Models I and II. For Model I, we get the estimate of the average variance in perceived quality evaluations across the four brands to be 0.148 and for Model II, we get the estimate of the average variance to be 0.651. This implies that even at the household level Model I predicts a smaller variance in perceived quality evaluations across the four brands as compared to Model II. In order to test the second implication, we compare the predicted extents of learning from consumption of a brand at the household level. Figure 2 shows the evolution of mean quality beliefs as predicted by Models I and II for a consumer who repeatedly purchases Surf for fifteen purchase occasions. Since the extent of learning from consumption is measured by total change in mean quality beliefs from the initial value of zero in Figure 2, we can infer that Model II predicts a higher extent of cumulative learning from consumption as compared to Model I.

In summary, a model that does not incorporate forgetting would tend to over-estimate the extent of consumer’s cumulative learning from past consumption experiences. As a result, such a

---

\(^{21}\) Note that the mean values of the perceived quality beliefs are not estimated in the model. They are computed by running the simulations over the observed choices of all consumers using the parameter estimates given in Table 2. This procedure is similar to the one use by Mehta, Rajiv and Srinivasan (2003)
model would make an incorrect prediction of the variance in the consumer’s perceived quality
evaluations of all brands.

**Consumer’s Learning from Consumption:** The estimate of $\alpha_s$ – the value of precision of
consumer’s quality beliefs at the beginning of the observation period in the estimation sample for
Model I is 2.11 and for Model II is 1.48. Recall that a high value of $\alpha_s$ implies that the extent of
initial quality learning that can occur through consumption is low. Since the estimate of $\alpha_s$ is
higher for Model I, it implies that the extent of learning from consumption at the beginning of the
observation period in the data is lower for Model I as compared to Model II. This highlights the
fact that if we do not explicitly take forgetting into account while measuring the extent of learning
in the data, we can overestimate the extent of initial learning through consumption. This is
because in Model II, the inter-temporal variation in a consumer’s overall quality evaluations is
attributed only to consumer’s learning from consumption experiences. On the other hand, the
inter-temporal variation in consumer’s quality evaluations in Model I is attributed both to
consumer’s learning from consumption and forgetting of prior evaluations.

Further since Model II does not incorporate forgetting, it would predict the extent of
learning from consumption to diminish with repeated consumption. On the other hand, since
Model I incorporates forgetting, it would predict uncertainty in quality evaluations (and hence the
learning from future consumption) to persist even after repeated consumption. This can be seen in
Figure 2: while Model II predicts the classic learning curve in which the learning from
consumption decreases with the number of repeat purchases, Model I predicts learning from
consumption to persist even after fifteen repeat purchases of the brand.

**Quality-Price Trade-off and Consumer Price Sensitivity:** The estimates of the mean quality
sensitivity parameter, $\bar{\theta}$, for Model I and II are 4.916 and 4.752 respectively. Even though this
might suggest that ignoring forgetting would lead to an over-estimation of price sensitivity (high
quality sensitivity or a high marginal willingness-to-pay for quality is synonymous to low price
sensitivity), we see in Tables 6A and 6B that the predicted price elasticities for Model I are much
higher than those for Model II. This is because the predicted difference in perceived quality
beliefs across the four brands is much smaller for the Model I, which results in a higher marginal
impact of price discounts in Model I. This highlights the fact that we can seriously under-estimate
the price elasticity if we do not take forgetting into account.
The estimate of the extent of heterogeneity in the quality sensitivity for Models I and II, \( \sigma_\theta \), is reported in Table 3. We get the estimate of \( \sigma_\theta \) from Model I and II as 0.521 and 1.100 respectively. This highlights the fact that ignoring forgetting may lead to overestimation of consumer heterogeneity in price sensitivities. This is because in Model II, all the variance in the purchase behavior across consumers is attributed to variance in price sensitivity and the heterogeneity in the extent of learning from consumption. On the other hand in Model I, the variance in the purchase behavior across consumers is attributed not only to variance in quality sensitivity and the heterogeneity in the extent of learning from consumption, but also to the heterogeneity in the extent of forgetting of prior quality evaluations.

**Determinants of Consumer’s Extent of Forgetting:** As discussed in section 2.6, if the consumer does not consume a brand in time periods following purchase occasion \( t \), her quality evaluations of that brand would decay to her prior quality evaluations at the beginning of her consumption history. Further, the extent of decay in her quality evaluations depends on (i) her forgetting parameter, \( b_i \), and (ii) the ratio of the information in the consumption set to the information in the positioning set, \( k_i = \frac{\alpha_i - \alpha_0}{\alpha_0} \). For Model I, we get the estimate of the consumer’s precision in her quality evaluations based on her positioning set as \( \alpha_0 = 0.636 \), the estimate of the mean value of \( b_i \) across the consumer population as \( \bar{b} = 0.0156 \) and the estimate of the extent of heterogeneity in the extent of forgetting as \( \sigma_b = 0.002 \).

The estimate for \( \bar{b} \) suggests that on an average it would take \( \frac{\ln(2)}{b} = 44 \) weeks for the information set corresponding to the information that the consumers had learned about a brand’s quality through prior consumption to depreciate to half of its value. This implies that the rate of decay of consumer’s quality evaluations of brands in the liquid detergent category is slow. This result is not surprising since experimental results in attitude persistence have shown that people’s overall evaluations formed through direct behavioral experiences remain fairly persistent across time and decay slowly (Fazio and Zanna 1981).

For Model I, we get the predicted average value of the precision in the consumer’s quality evaluations (averaged across all purchase observation and all brands in the data) to be \( \alpha_{t,avg} = 4.1 \). It follows that the average value of the ratio of the information in the consumer’s consumption set to the information in her positioning set would be \( k_{t,avg} = 5.45 \). Since \( k_{t,avg} > 1 \), it implies that on average, the consumer’s quality evaluations would decay in a sigmoid fashion in absence of
consumption. This can be seen in Figure 3 where we plot the decay of consumer’s mean quality evaluations of a brand with time in absence if its consumption for the estimated values of $\bar{b}$, $\alpha_0$ and the average value of the level of precision of consumer’s quality evaluations in the data, $\alpha_{t,\text{avg}} = 4.1$. In order to demonstrate the extent of heterogeneity in the parameter $b_i$ across the consumer population, we also plot the decay of mean quality evaluations for two other values of $b_i$ that are one standard deviation away from the mean value, that is, $b_i = \bar{b} + \sigma_b$ and $b_i = \bar{b} - \sigma_b$.

Observe that the quality evaluations for the three values of $b_i$ show a significant difference only in the later part of the graph. This implies that in order to estimate the heterogeneity significantly, we would need to estimate the proposed model on a data set that covers a long period of time – which explains why we do not get very significant estimates of the extent of heterogeneity in forgetting since our scanner data only extends for 111 weeks.

**Impact of Forgetting on Habit Persistence and State Dependence:** The decrease in habit persistence with the length of inter-purchase time is similar to the decay of the mean quality evaluations. Figure 4 plots the decrease in habit persistence versus the inter-purchase time as predicted by Model I. The plot is shown using the estimated values of $\bar{b}$, $\alpha_0$ and the average value of precision of quality evaluations in the data, $\alpha_{t,\text{avg}} = 4.1$. Observe that the decay of habit persistence is rather modest since the extent of forgetting, $\bar{b}$, is small.

Figure 5 shows the plot of last period’s state dependence versus the length of the inter-purchase time as predicted by Model I and SAC (1999). In order to facilitate this comparison, we plot the ratio of the state dependence in consumer’s utility for a given inter-purchase time to the state dependence in consumer’s utility if her inter-purchase time, $w_{t,t+1}$, were zero. For Model I, we plot the state dependence for the estimated values of $\bar{b}$, $\alpha_0$ and the average value of precision of quality evaluations in the data. Note that the increase in state dependence with time as predicted by the SAC (1999) is higher than that predicted by Model I. This is because the SAC (1999) does not take the decrease of habit persistence with time into account and as a result it leads to a biased prediction of the change in consumer’s state dependence.

4. CONCLUSIONS

In this paper, we propose a structural model of forgetting in which the consumer imperfectly recalls her prior overall quality evaluations while constructing her new evaluations on a purchase occasion. We conceptualize the imperfect recall by positing that the consumer recalls
her prior quality evaluations with an added noise that results in her recalled quality evaluations to
be different from her prior evaluations. The extent of the added noise is characterized as a
function of the time lapsed since the prior purchase occasion suggesting that the extent of
forgetting is an increasing and concave function of time. Our main analytical results are: (i) In
absence of consumption, the expected value of consumer’s quality evaluation decays over time in
a sigmoid fashion if the information in her consumption set is greater than the information in her
positioning set. On the other hand, the consumer’s expected quality evaluation decays in a convex
fashion if the information in her consumption set is less than the information in her positioning
set. (ii) The greater the information in consumer’s consumption set, the smaller is the decay of her
mean quality evaluations over time. (iii) The heterogeneity in both habit persistence and state
dependence results from heterogeneity in consumers’ willingness to pay for quality and the
heterogeneity in their extent of forgetting of prior evaluations. (iv) The habit persistence in
consumer’s utility decreases with the length of the inter-purchase time as a result of forgetting.
Further, the greater the information in the consumer’s consumption set, the smaller is the decrease
in habit persistence over time. (v) The last period’s state dependence in consumer’s utility
increases with the length of the inter-purchase time. And the greater the information in the
consumer’s consumption set, the smaller is the increase in state dependence over time.

In our empirical analysis on the liquid detergent data set, we find that on an average a
small, but significant forgetting of prior quality evaluations per week. We find that ignoring
forgetting in choice models might lead to (i) an over estimation of the extent of initial learning
that can be done through consumption and under prediction of the learning that can be done
through consumption in later stages of the consumer’s consumption history; (ii) an over
estimation of the variance in perceived quality beliefs across the given brands that leads to an
over prediction of the price elasticities; (iii) an over estimation of the heterogeneity in the price
sensitivity across the consumer population.

We consider this paper as an important first step to study effect of forgetting in the context
of frequently purchased products 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, we do not consider the impact of forgetting on brand recognition that might
impact the consumer’s consideration set formation (Nedungudi 1990). Second because of lack of
data on consumers’ exposure to TV ads, we do not consider the forgetting of quality evaluations
formed through TV ads. Prior research suggests that decay of quality evaluations formed from
indirect experiences like TV ads would be greater than the decay of quality evaluations formed
from direct consumption experiences (Fazio and Zanna 1981).
TABLES AND FIGURES

TABLE 1: Descriptive Statistics for Liquid Detergent Data Set

<table>
<thead>
<tr>
<th>Brand</th>
<th>Market Share</th>
<th>Mean Price (Std. Dev.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wisk</td>
<td>0.31</td>
<td>5.06 (0.58)</td>
</tr>
<tr>
<td>Tide</td>
<td>0.28</td>
<td>5.97 (0.61)</td>
</tr>
<tr>
<td>Era</td>
<td>0.21</td>
<td>5.92 (0.57)</td>
</tr>
<tr>
<td>Surf</td>
<td>0.20</td>
<td>5.46 (0.61)</td>
</tr>
</tbody>
</table>

NOTE: Price is in cents per ounce.

Exhibit 1: Distribution of Inter-Purchase Times across the Consumer Population for Liquid Detergent Data Set

Distribution of Inter-Purchase Times for Liquid Detergents

![Graph showing distribution of inter-purchase times in weeks with frequency across the population on the y-axis and inter-purchase time in weeks on the x-axis.]
### TABLE 2: Posterior Quality Beliefs of Brands from the Pre-Estimation Sample for Liquid Detergent Data Set

<table>
<thead>
<tr>
<th>Mean of the Posterior Density</th>
<th>MODEL I</th>
<th>MODEL II</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega_{\text{Wisk}}$ (Mean of the Posterior Quality Belief of Wisk)</td>
<td>0.0089</td>
<td>0.0226</td>
</tr>
<tr>
<td>$\omega_{\text{Tide}}$ (Mean of the Posterior Quality Belief of Tide)</td>
<td>0.0192</td>
<td>0.0835</td>
</tr>
<tr>
<td>$\omega_{\text{Era}}$ (Mean of the Posterior Quality Belief of Era)</td>
<td>0.0101</td>
<td>0.0430</td>
</tr>
<tr>
<td>$\omega_{\text{Surf}}$ (Mean of the Posterior Quality belief of Surf)</td>
<td>0.0076</td>
<td>0.0261</td>
</tr>
</tbody>
</table>

### TABLE 3: Parameter Estimates for Models I & II for Liquid Detergent Data Set

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Explanation</th>
<th>MODEL I (Std. Deviation)</th>
<th>MODEL II (Std. Deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{\theta}$</td>
<td>Mean value of the intensity of preference for quality across the consumers</td>
<td>4.916 (0.280)</td>
<td>4.752 (0.141)</td>
</tr>
<tr>
<td>$\sigma_{\theta}$</td>
<td>Standard Deviation of the distribution of intensity of preference for quality across the consumers</td>
<td>0.521 (0.220)</td>
<td>1.100 (0.397)</td>
</tr>
<tr>
<td>$\alpha_{s}$</td>
<td>Inverse of the uncertainty in quality of a brand at the beginning of the observation period in the data; assumed same $\forall i, \forall j$</td>
<td>2.110 (0.461)</td>
<td>1.480 (0.671)</td>
</tr>
<tr>
<td>$\alpha_{0}$</td>
<td>Inverse of the uncertainty in quality of a brand at the beginning of their consumption history; assumed same $\forall i, \forall j$</td>
<td>0.636 (0.080)</td>
<td>-</td>
</tr>
<tr>
<td>$\bar{b}$</td>
<td>Mean value of the log of decay of the information learned from past consumption experiences due to forgetting across the consumer population</td>
<td>0.0156 (0.002)</td>
<td>-</td>
</tr>
<tr>
<td>$\sigma_{b}$</td>
<td>Standard Deviation of the distribution of the log of decay of the information learned from past consumption experiences due to forgetting across the consumer population</td>
<td>0.002 (0.0012)</td>
<td>-</td>
</tr>
<tr>
<td>$\omega_{\text{Wisk}}$</td>
<td>Predicted Mean value of perceived quality of Wisk (across all consumers and across all purchase occasions)</td>
<td>0.0110</td>
<td>0.0431</td>
</tr>
</tbody>
</table>

---

22. 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>
<thead>
<tr>
<th>Parameter</th>
<th>Explanation</th>
<th>MODEL I (Std. Deviation)</th>
<th>MODEL II (Std. Deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega_{\text{Tide}}$</td>
<td>Predicted Mean value of perceived quality of Tide (across all consumers and across all purchase occasions)</td>
<td>0.0203</td>
<td>0.0624</td>
</tr>
<tr>
<td>$\omega_{\text{Era}}$</td>
<td>Predicted Mean value of perceived quality of Era (across all consumers and across all purchase occasions)</td>
<td>0.0103</td>
<td>0.0469</td>
</tr>
<tr>
<td>$\omega_{\text{Surf}}$</td>
<td>Predicted Mean value of perceived quality of Surf (across all consumers and across all purchase occasions)</td>
<td>0.0091</td>
<td>0.0307</td>
</tr>
</tbody>
</table>

**TABLE 4: Parameter Estimates for SAC (1999) for Liquid Detergent Data Set**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Explanation</th>
<th>Estimates (Std. Deviation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega_{\text{Tide}}$</td>
<td>Intercept for Tide ($\omega_{n,k} = 0$ for identification)</td>
<td>1.436</td>
</tr>
<tr>
<td>$\omega_{\text{Era}}$</td>
<td>Intercept for Era</td>
<td>0.818</td>
</tr>
<tr>
<td>$\omega_{\text{Surf}}$</td>
<td>Intercept for Surf</td>
<td>0.309</td>
</tr>
<tr>
<td>$\bar{\beta}_{\text{price}}$</td>
<td>Mean value of the coefficient of the log of price in the utility across the consumer population</td>
<td>7.381</td>
</tr>
<tr>
<td>$\bar{\beta}_{sd}$</td>
<td>Mean value of the coefficient of state dependence in the utility across the consumer population</td>
<td>1.216</td>
</tr>
<tr>
<td>$\bar{\beta}_{wo}$</td>
<td>Mean value of the coefficient of wear out effect of state dependence in the utility across the consumer population</td>
<td>0.186</td>
</tr>
<tr>
<td>$\sigma_{\omega_{\text{Tide}}}$</td>
<td>Standard Deviation of the distribution of intercept of Tide across the consumer population</td>
<td>0.752</td>
</tr>
<tr>
<td>$\sigma_{\omega_{\text{Era}}}$</td>
<td>Standard Deviation of the intercept of Era across the consumer population</td>
<td>0.922</td>
</tr>
<tr>
<td>$\sigma_{\omega_{\text{Surf}}}$</td>
<td>Standard Deviation of the distribution of the intercept of Surf across the consumer population</td>
<td>0.706</td>
</tr>
<tr>
<td>$\sigma_{\beta_{\text{price}}}$</td>
<td>Standard Deviation of the distribution of the coefficient of the log of price in the utility across the consumer population</td>
<td>2.990</td>
</tr>
<tr>
<td>$\sigma_{\beta_{sd}}$</td>
<td>Standard Deviation of the coefficient of state dependence in the utility across the consumer population</td>
<td>0.330</td>
</tr>
<tr>
<td>$\sigma_{\beta_{wo}}$</td>
<td>Standard Deviation of the distribution of the coefficient of wear out effect of state dependence in the utility across the consumer population</td>
<td>0.106</td>
</tr>
</tbody>
</table>
TABLE 5A: Result from comparison of Proposed Model (MODEL I) with Competing Models for Liquid Detergents (MODEL II, SAC (1999))

<table>
<thead>
<tr>
<th>Model Comparison using Singleton Test</th>
<th>Singleton Test Results</th>
<th></th>
</tr>
</thead>
</table>
| MODEL I against MODEL II             | \( J_S = 5.72 (\chi^2 \text{ with 1 d.f.}) \)  
\( p\)-value < 0.05                  | \( J_S = 6.14 (\chi^2 \text{ with 1 d.f.}) \)  
\( p\)-value < 0.02                  |  |
| MODEL I against SAC (1999)           | \( J_S = 3.10 (\chi^2 \text{ with 1 d.f.}) \)  
\( p\)-value < 0.10                  | \( J_S = 5.55 (\chi^2 \text{ with 1 d.f.}) \)  
\( p\)-value < 0.05                  |  |

TABLE 5B: Hit rates and Likelihood values for Models I, II, SAC (1999) for Hold out Sample for Liquid Detergents

<table>
<thead>
<tr>
<th></th>
<th>MODEL I</th>
<th>MODEL II</th>
<th>SAC (1999)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hit Rate</td>
<td>69.12%</td>
<td>63.33%</td>
<td>64.80%</td>
</tr>
<tr>
<td>-Log Likelihood</td>
<td>2066</td>
<td>2478</td>
<td>2251</td>
</tr>
</tbody>
</table>

TABLE 6A: Model I – Price Elasticities for Liquid Detergents

<table>
<thead>
<tr>
<th>Price Elasticity</th>
<th>Wisk Market Share</th>
<th>Tide Market Share</th>
<th>Era Market Share</th>
<th>Surf Market Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price of Wisk</td>
<td>-2.466</td>
<td>1.045</td>
<td>1.164</td>
<td>1.470</td>
</tr>
<tr>
<td>Price of Tide</td>
<td>0.740</td>
<td>-3.338</td>
<td>0.721</td>
<td>0.818</td>
</tr>
<tr>
<td>Price of Era</td>
<td>0.720</td>
<td>0.591</td>
<td>-2.910</td>
<td>0.746</td>
</tr>
<tr>
<td>Price of Surf</td>
<td>0.851</td>
<td>0.680</td>
<td>0.752</td>
<td>-2.660</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Price Elasticity</th>
<th>Wisk Market Share</th>
<th>Tide Market Share</th>
<th>Era Market Share</th>
<th>Surf Market Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price of Wisk</td>
<td>-1.285</td>
<td>0.439</td>
<td>0.574</td>
<td>1.167</td>
</tr>
<tr>
<td>Price of Tide</td>
<td>0.272</td>
<td>-1.231</td>
<td>0.431</td>
<td>0.462</td>
</tr>
<tr>
<td>Price of Era</td>
<td>0.267</td>
<td>0.020</td>
<td>-1.430</td>
<td>0.352</td>
</tr>
<tr>
<td>Price of Surf</td>
<td>0.571</td>
<td>0.418</td>
<td>0.459</td>
<td>-1.490</td>
</tr>
</tbody>
</table>
Figure 1A: Decay of Mean Quality Beliefs with Time in absence of Consumption for

\[ k_{it} = \frac{\alpha_i - \alpha_0}{\alpha_0} = 10 \]

Figure 1B: Decay of Mean Quality Beliefs with Time in absence of Consumption for Extent of Forgetting, \( b_i = 0.015 \)
Figure 2: The Evolution of Mean Quality Beliefs of Surf for 15 Purchase Occasions for a Surf Loyal Consumer as Predicted by Model I and Model II

![Evolution of Mean Quality Beliefs](image)

Figure 3: Decay of Mean Quality Beliefs in absence of Consumption since Purchase Occasion \( t \) as predicted by Model I for Liquid Detergents

![Decay of Mean Quality Beliefs](image)
Figure 4: Decay of Habit Persistence with the Inter-purchase Time (between purchase occasions $t$ and $t+1$)

![Graph showing habit persistence decay over time.]

Initial Precision = 4.1

Figure 5: State Dependence vs. Inter-purchase time (between purchase occasions $t$ and $t+1$) as predicted by Model I and SAC (1999)

![Graph showing state dependence over time.]

Model I for Initial Precision = 4.1

SAC (1999)
REFERENCES


APPENDIX

Proposition 1: As long as $b_i > 0$, $\sigma^2_q > 0$, and $w_{i,r,t} > 0 \forall t$, consumer $i$ can never learn about her true quality evaluation of a brand even after repeat purchasing it infinite times.

Proof: Consider a consumer $i$ whose overall evaluations about brand $j$’s quality at purchase occasion $t-1$ are $q_{ij,t-1} \sim N(\omega_{ij,t-1}, \sigma^2_{ij,t-1})$ and purchases brand $j$ on every purchase occasion thereafter. Represent the shortest purchase interval after purchase occasion $t-1$ as ‘$w$’ such that $w = \min \{w_{i,r,t-1} \} \forall r$. In the following analysis, we will prove that even if all the purchase intervals after purchase occasion $t-1$ were of length ‘$w$’, the consumer’s evaluations as purchase occasion $r$ as $r \to \infty$ would still not converge to her true quality evaluation $q_{ij}$. This would imply that for any other distribution of inter-purchase times from purchase occasion $t-1$ till $\infty$, the consumer’s overall evaluations would not converge to her true quality evaluation $q_{ij}$.

Since consumer $i$ purchases brand $j$ on every purchase occasion after $t-1$, represent the quality signals that she received from consumption of brand $j$ by $\{ \hat{q}_{ij,r} \} \forall r \geq t$. Using equation (10) in the paper, the mean and precision of the overall quality evaluations of brand $j$ at purchase occasion $r$ can be written in terms of the mean and precision of the overall quality evaluations at purchase occasion $r-1$ as follows:

$$\omega_{ij,r} = \frac{f\alpha_{ij,r-1}\omega_{ij,r-1} + (1-f)\omega_{ij,0}\alpha_{ij,0} + \hat{v}_{ij,r-1}\sqrt{\alpha_{ij,r-1} - \alpha_{ij,0}}f(1-f) + \hat{\lambda}_{ij,r-1}}{f\alpha_{ij,r-1} + (1-f)\alpha_{ij,0} + 1}$$

(A1)

$$a_{ij,r} = f\alpha_{ij,r-1} + (1-f)\alpha_{ij,0} + 1$$

where $r > t-1$ and $f = \exp(-b_i w)$. Note that the consumptions signal at any purchase occasion $r-1$ is distributed as $\lambda_{ij,r-1} \sim N(\hat{q}_{ij}, \sigma^2_q)$, is i.i.d across all purchase occasions and is uncorrelated with the forgetting noise $v_{ij,r-1}$. Substituting $\hat{q}_{ij,r-1} \sim N(\hat{q}_{ij}, \sigma^2_q)$ into equation (A1), we get the expression for the mean quality evaluation of brand $j$ at occasion $r$ as follows

$$\omega_{ij,r} = \frac{f\alpha_{ij,r-1}\omega_{ij,r-1} + (1-f)\omega_{ij,0}\alpha_{ij,0} + \hat{\epsilon}_{ij,r-1}\sqrt{\alpha_{ij,r-1} - \alpha_{ij,0}}f(1-f) + \hat{\lambda}_{ij,r-1}}{f\alpha_{ij,r-1} + (1-f)\alpha_{ij,0} + 1}$$

(A2)

where $\hat{\epsilon}_{ij,r-1}$ is a normal random variable that represents the total noise due to learning from the consumption signal $\hat{\lambda}_{ij,r-1}$ and the forgetting of quality evaluations at purchase occasion $r-1$. Further note that $\hat{\epsilon}_{ij,r-1}$ is independent across all consumers $i$, brands $j$ and purchase occasions $r$. Using equations (A1) and (A2), we can recursively derive the consumer $i$’s evaluations of brand $j$’s quality at purchase occasion $r$ as a function of her beliefs of brand $j$’s quality at purchase occasion $t-1$ as follows

$$a_{ij,r} = a_{ij,r-1} + \frac{f - \alpha_{ij,r-1}}{1-f} + \alpha_{ij,0}(1-f)$$

(A3)

$$\omega_{ij,r} = \frac{1}{f^{r+1}a_{ij,r-1} + (1-f^{r+1})\alpha_{ij,0} + 1 - f^{r+1}} \left[ \frac{f^{r+1}a_{ij,r-1}\omega_{ij,r-1} + (1-f^{r+1})\omega_{ij,0}\alpha_{ij,0} + \hat{\epsilon}_{ij,r-1}\sqrt{\alpha_{ij,r-1} - \alpha_{ij,0}}f(1-f) + \hat{\lambda}_{ij,r-1}}{f^{r+1}\alpha_{ij,r-1} + (1-f^{r+1})\alpha_{ij,0} + 1} \right] + \hat{\lambda}_{ij,r-1}$$
where \( \hat{\epsilon}_{ij,r-1} \left[ 1 - f^{r-1} \right] \left[ 1 \left( a_{ij,r-1} - a_{ij,0} \right) f^{r-1} \left( 1 - f \right) \right] \) is a normal random variable that represents the cumulative noise due to learning and forgetting of evaluations of brand \( j \) from purchase occasion \( t-1 \) till \( r \).

Observe that in the limit as \( r \) approaches infinity, the mean and precision of quality evaluations converge to

\[
\lim_{r \to \infty} \alpha_{ij,r} = \frac{1}{1 - f} + \alpha_{ij,0}
\]

\[
\lim_{r \to \infty} \omega_{ij,r} = \frac{q_{ij} + \left( 1 - f \right) \alpha_{ij,0} \omega_{ij,0} + \hat{\epsilon}_{ij,\infty,t-1}}{1 + \left( 1 - f \right) \alpha_{ij,0}} + \frac{\sqrt{1 - f}}{1 + \left( 1 - f \right) \alpha_{ij,0}}
\]

(A4)

This proves that even after infinite consumption signals, a consumer would still be uncertain about her true quality evaluation of the brand as long as \( f \) is strictly less than one (that is, \( w > 0 \) and \( b > 0 \)).

Proposition 2: In absence of purchase of a brand at purchase occasion \( t \) and thereafter, the consumer’s mean quality evaluation of that brand reverses back with time to her mean quality evaluation at the beginning of her purchase history. Further,

(i) The mean quality evaluations decay in a sigmoid fashion if the information in the consumer’s consumption set, \( \alpha_{ij} - \alpha_0 \), is greater than the information in her positioning set, \( \alpha_0 \).

(ii) The mean quality evaluations decay in a convex fashion if the information in the consumer’s consumption set, \( \alpha_{ij} - \alpha_0 \), is smaller than the information in her positioning set, \( \alpha_0 \).

(iii) The extent of reversion of the mean quality evaluation over time is an increasing function of consumer’s extent of forgetting per week, \( b_i \).

(iv) The extent of reversion of the mean quality evaluation is a decreasing function of the ratio of the consumption set, to the information in her positioning set, \( \alpha_{ij} \).

Proof: Consider a consumer \( i \) whose evaluations about brand \( j \)’s quality at purchase occasion \( t-1 \) are \( q_{ij,t-1} \sim N\left( \alpha_{ij,t-1}, \sigma_{ij,t-1}^2 \right) \). Further assume that she does not purchase brand \( j \) after purchase occasion \( t-1 \).

Using equation (10) in the paper, we can write her quality evaluations on purchase occasion \( r \) (where \( r > t-1 \)) as a function of her evaluations at occasion \( r-1 \) as follows

\[
\omega_{ij,r} = \frac{e^{-b_{ij,r-1}} \alpha_{ij,r-1} a_{ij,r-1} a_{ij,0} + \left( 1 - e^{-b_{ij,r-1}} \right) a_{ij,0} \alpha_{ij,0} + \hat{\epsilon}_{ij,r-1} / \alpha_{ij,0} \left( 1 - e^{-b_{ij,r-1}} \right)}{e^{-b_{ij,r-1}} \alpha_{ij,r-1} + \left( 1 - e^{-b_{ij,r-1}} \right) \alpha_{ij,0}}
\]

(A5)

Since the random error \( v_{ij,r-1} \) is independent across all \( r \), we can recursively derive the mean quality evaluations of consumer \( i \) on purchase occasion \( r \) as a function of her evaluations at occasion \( t-1 \) as follows

\[
\omega_{ij,r} = \frac{e^{-b_{ij,r-1}} \alpha_{ij,r-1} a_{ij,r-1} a_{ij,0} + \left( 1 - e^{-b_{ij,r-1}} \right) a_{ij,0} \alpha_{ij,0} + \hat{\epsilon}_{ij,r-1} / \alpha_{ij,0} \left( 1 - e^{-b_{ij,r-1}} \right)}{e^{-b_{ij,r-1}} \alpha_{ij,r-1} + \left( 1 - e^{-b_{ij,r-1}} \right) \alpha_{ij,0}}
\]

(A6)

where \( w_{ij,r-1} \) is the time lapsed between purchase occasions \( t-1 \) and \( r \). The first two terms in the RHS of equation (A6) represent the deterministic components of the mean quality evaluation of brand \( j \) at occasion \( r \) and the third term in the RHS of equation (A6) is a normal random variable that represents the cumulative noise that gets added from purchase occasion \( t-1 \) till \( r \) due to forgetting. In order to prove proposition 2, we will only deal with the deterministic component of the mean quality evaluation at time \( r \). Writing equation (A6) in terms of the expected difference between the mean quality evaluation at occasion \( r \), \( \omega_{ij,r} \), and the mean evaluation at the beginning of the consumer’s purchase history, \( \omega_{ij,0} \), (conditional on her information set at purchase occasion \( t-1 \)) we get
\[
\omega_{ij,t}^d = \omega_{ij,t-1}^d \frac{e^{-b_{ij,t-1,r}} \alpha_{ij,t-1}}{e^{-b_{ij,t-1,r}} (\alpha_{ij,t-1} - \alpha_{ij,0}) + \alpha_{ij,0}}. \tag{A7}
\]

where \(\omega_{ij,t}^d = E(\omega_{ij,t} - \omega_{ij,0}) \forall r\). We can represent equation (A7) in terms of the ratio of the information that the consumer had learned from her consumption experiences of that brand prior to occasion \(t-1\) to the information she had about the quality of that brand at the beginning of her purchase history, that is,

\[
k_{i,t-1} = \frac{\alpha_{ij,t-1} - \alpha_{ij,0}}{\alpha_{ij,0}}, \text{ as follows}
\]

\[
\omega_{ij,t}^d = \omega_{ij,t-1}^d \frac{e^{-b_{ij,t-1,r}} (k_{i,t-1} + 1)}{e^{-b_{ij,t-1,r}} k_{i,t-1} + 1}. \tag{A8}
\]

Taking the first derivative of \(\omega_{ij,t}^d\) with respect to the time lapsed between purchase occasion \(t-1\) and \(r\), we get

\[
\frac{\partial \omega_{ij,t}^d}{\partial w_{i,t-1,r}} = -\omega_{ij,t-1}^d b_i e^{b_{ij,t-1,r}} (k_{i,t-1} + 1) \left( e^{b_{ij,t-1,r}} k_{i,t-1} + 1 \right)^{-2}
\]

Note that the RHS of equation (A9) is negative. This implies that the mean quality evaluations decay to the prior mean quality evaluations with time. Taking the second derivative of \(\omega_{ij,t}^d\) with respect to \(w_{i,t-1,r}\), we get

\[
\frac{\partial^2 \omega_{ij,t}^d}{\partial w_{i,t-1,r}^2} = \omega_{ij,t-1}^d \left( e^{b_{ij,t-1,r}} - k_{i,t-1} \right) b_i^2 e^{b_{ij,t-1,r}} (k_{i,t-1} + 1) \left( e^{b_{ij,t-1,r}} k_{i,t-1} + 1 \right)^{-3}
\]

Note that if \(k_{i,t-1} < 1\), the second derivative will be positive. This implies that as long as \(\alpha_{ij,t-1} - \alpha_{ij,0} < \alpha_{ij,0}\), the evaluations will decay in a convex fashion. Further if, \(k_{i,t-1} > 1\), note the following: the second derivative is less than zero as long as the time lapsed since purchase occasion \(t-1\), \(w_{i,t-1,r} < \frac{\ln(k_{i,t-1})}{b_i}\) and is greater than zero when the time lapsed since occasion \(t-1\), \(w_{i,t-1,r} > \frac{\ln(k_{i,t-1})}{b_i}\).

This shows that the mean quality evaluations decay in a sigmoid fashion when \(\alpha_{ij,t-1} - \alpha_{ij,0} > \alpha_{ij,0}\) where the rate of decay increases till time \(\frac{\ln(k_{i,t-1})}{b_i}\) and decreases after that.

In order to prove that the extent of decay of mean quality evaluations with time is a decreasing function of \(k_{i,t-1}\), we take the first derivative of \(\omega_{ij,t}^d\) with respect to \(k_{i,t-1}\) to get

\[
\frac{\partial \omega_{ij,t}^d}{\partial k_{i,t-1}} = \omega_{ij,t-1}^d \frac{e^{b_{ij,t-1,r}} - 1}{e^{b_{ij,t-1,r}} k_{i,t-1} + 1}^2
\]

Note that the RHS of equation (A11) is positive. Since the extent of decay of mean quality evaluations is given by \(\omega_{ij,t-1}^d - \omega_{ij,t}^d\), this proves the third part of the proposition that the extent of decay is a decreasing function of \(k_{i,t-1}\).

**Proposition 3:** The habit persistence in consumer’s purchase behavior at purchase occasion \(t\) decreases with the length of the inter-purchase time between occasions \(t-1\) and \(t\) as a result of forgetting. Further, the extent of decay of the habit persistence with the length of the inter-purchase time is a decreasing function of ratio of the information that the consumer had learned from her consumption experiences prior to occasion \(t-1\) to the information she had learned about the quality of that brand at the beginning of her purchase history.

**Proof:** As discussed in Section 2.33 of the paper, the habit persistence in consumer \(i\)’s utility for brand \(j\) at purchase occasion \(t\) as follows

\[
41
\]
Proof:  

\[ h_i = \frac{\theta_i \alpha_i w_{i,t-1} e^{h w_{i,t-1}}}{e^{h w_{i,t-1}} \alpha_{i,t-1} + (1 - e^{h w_{i,t-1}}) \alpha_{i,0} + d_{i,t-1}} \]  \tag{A12}\]

Next, we write equation (A12) in terms of ratio of the information that the consumer had learned from her consumption experiences prior to occasion \( t-1 \) to the information she had learned about the quality of that brand at the beginning of her purchase history, that is \( k_{i,t-1} = \frac{\alpha_{i,t-1} - \alpha_{i,0}}{\alpha_{i,0}} \), as follows

\[ h_i = \frac{\theta_i \alpha_i w_{i,t-1} (k_{i,t-1} + 1)}{1 + e^{k_{i,t-1}} k_{i,t-1} \alpha_{i,0} + d_{i,t-1}} \]  \tag{A13}\]

Taking the first derivative of \( h_i \) with respect to \( w_{i,t-1,t} \), we get the following

\[ \frac{\partial h_i}{\partial w_{i,t-1,t}} = \frac{\theta_i \alpha_i w_{i,t-1} (k_{i,t-1} + 1) + (\alpha_i + d_{i,t-1})}{(1 + e^{k_{i,t-1}} (\alpha_{i,0} + d_{i,t-1} + \alpha_i k_{i,t-1}))} \]  \tag{A14}\]

Note that the RHS of equation (A14) is negative. This proves the first part of the proposition that habit persistence at purchase occasion \( t \) decreases with the length of prior inter-purchase time \( w_{i,t-1,t} \). In order to prove the second part of the proposition that the extent of decay of the habit persistence with the length of the inter-purchase time is a decreasing function of ratio of the information that the consumer had learned from her consumption experiences prior to occasion \( t-1 \) to the information she had learned about the quality of that brand at the beginning of her purchase history, that is, \( k_{i,t-1} \), we take the first derivative of \( h_i \) with respect to \( k_{i,t-1} \) to get the following

\[ \frac{\partial h_i}{\partial k_{i,t-1}} = \frac{\theta_i \alpha_i w_{i,t-1} (k_{i,t-1} + 1) + (\alpha_i + d_{i,t-1})}{(1 + e^{k_{i,t-1}} (\alpha_{i,0} + d_{i,t-1} + \alpha_i k_{i,t-1}))} \]  \tag{A15}\]

Note that the RHS of the equation (A15) is positive. Since the extent of decay of habit persistence is given by \( h_0 - h_i \), this proves that the extent of decay of habit persistence is a decreasing function of \( k_{i,t-1} \).

**Proposition 4:** Consumer’s state dependence in her purchase behavior at purchase occasion \( t \) increases with the length of her inter-purchase time between occasions \( t-1 \) and \( t \) as a result of forgetting. Further, the extent of increase of the state dependence with the length of the inter-purchase time is a decreasing function of ratio of the information that the consumer had learned from her consumption experiences prior to occasion \( t-1 \) to the information she had about the quality of that brand at the beginning of her purchase history.

**Proof:** As discussed in Section 2.52 of the paper, the term \( s_i = \frac{\theta_i \lambda_{i,t-1}}{e^{h w_{i,t-1}} \alpha_{i,t-1} + (1 - e^{h w_{i,t-1}}) \alpha_{i,0} + 1} \) represents the state dependence term in consumer \( i \)'s utility of brand \( j \) at purchase occasion \( t \).

Writing the expression for state dependence in terms of \( k_{i,t-1} = \frac{\alpha_{i,t-1} - \alpha_{i,0}}{\alpha_{i,0}} \), we get

\[ s_i = \frac{\theta_i \lambda_{i,t-1}}{1 + (1 + e^{k_{i,t-1}}) k_{i,t-1} \alpha_{i,0}} \]  \tag{A16}\]

Taking the first derivative of \( s_i \) with respect to \( w_{i,t-1,t} \), we get

\[ \frac{\partial s_i}{\partial w_{i,t-1,t}} = \frac{\theta_i \lambda_{i,t-1} \alpha_i w_{i,t-1} e^{h w_{i,t-1}}}{(\alpha_i + 1) e^{h w_{i,t-1}} + \alpha_i k_{i,t-1}} \]  \tag{A17}\]

Note that the RHS of equation (A17) is positive. This proves the first part of proposition 4 that state dependence at purchase occasion \( t \) decreases with the prior inter-purchase time, \( w_{i,t-1,t} \). In order to prove the second part of proposition 4 that the extent of increase of the habit persistence with the length of the inter-purchase time is a decreasing function of ratio of the information that the consumer had learned from her consumption experiences prior to occasion \( t-1 \) to the information she had learned about the quality of that brand at the beginning of her purchase history, that is, \( k_{i,t-1} \), we take the first derivative of \( s_i \) with respect to \( k_{i,t-1} \) to get
\[
\frac{\partial \tilde{S}_i}{\partial k_{i,j-1}} = -\frac{\partial \hat{\lambda}_{i,j-1} \alpha_{ij} e^{b_{i,j-1} j}}{(\alpha_0 + 1) e^{b_{i,j-1} j} + \alpha_0 k_{i,j-1}}
\]

Note that the RHS of the equation (A18) is negative. This proves that the extent of increase of state dependence is a decreasing function of \( k_{i,j-1} \).