Consumer Learning and Evolution of Consumer Brand Preferences

Hai Che*
Indiana University in Bloomington

Tülin Erdem
New York University

T.Sabri Öncü
Centre for Advanced Financial Research and Learning, Reserve Bank of India

*Corresponding author. Hai Che is Associate Professor of Marketing, Kelley School of Business (haiche@indiana.edu; (812) 855-4396). Tülin Erdem is Leonard N. Stern Professor of Business at Stern School of Business, New York University (terdem@stern.nyu.edu). T. Sabri Öncü is Head of Research at the Centre for Advanced Financial Research and Learning, Reserve Bank of India and Visiting Scholar at Stern School of Business, New York University (sabri.oncu@gmail.com). Authors are listed alphabetically and contribute equally to the paper.

Consumer Learning and Evolution of Consumer Brand Preferences

Abstract

We develop a structural dynamic demand model that examines how brand preferences evolve when consumers are new to a market and their needs change periodically. We allow for strategic sampling behavior of consumers under quality uncertainty. However, we differ from previous work on forward-looking consumer Bayesian learning by allowing for 1) spill-over learning effects across the sizes of each brand, 2) duration-dependence in utility for a brand-size to capture systematic periodic changes in consumer utility for a brand-size, 3) evolution of price sensitivities in markets where there is consumer quality uncertainty that diminishes over time as consumers get more experienced. We estimate our model using scanner data for the disposable diapers category and discuss the consumer behavior and managerial implications of our estimation and policy simulation results.

Keywords: Strategic Sampling, Spill-over effects, Duration Dependence, Consumer Choice under Uncertainty, Bayesian Learning

I. Introduction and Background

Previous literature on forward-looking Bayesian learning models has shown that in frequently purchased product categories, consumers may sample brands strategically; that is, they may forgo current utility to get information about brand quality and maximize expected utility over the planning horizon (e.g., Erdem and Keane 1996; Ackerberg 2003; Crawford and Shum 2005; Sun 2005; Hartmann 2006). Such dynamic structural demand models predict that consumers switch across brands relatively early on (due to strategic sampling) and later they settle on a small sub-set of brands once uncertainty is mostly resolved, implying that one observes more switching early on and less switching in later periods.

However, in markets where there is quality uncertainty, price sensitivity may change over time too.[1] If price sensitivities increase with reduced quality uncertainty, diminished brand switching due to reduced strategic sampling as consumers gain more experience with brands would be dampened by increased switching due to increased price sensitivities in later periods. Thus, overall switching may not decline (or decline less than what a forward-looking model with fixed price sensitivities would suggest) but the reasons for switching may change: early on, consumers may switch more for learning purposes whereas in later periods they may switch more due to price as they get more price sensitive. Indeed, when we inspect the brand switching patterns in the scanner panel purchase data for disposable diapers, we do not observe a declining trend in brand switching. To capture both consumer learning and changing price sensitivities as consumers become more experienced, it is important to study the behavior of new consumers to a market and observe their behavior over time.

[1] Heilman, Bowman and Wright's (2000) reduced-form model results indicated that price sensitivities decrease with increased use experience in disposable diapers category. They did not offer an explanation for this result.

Furthermore, standard forward-looking Bayesian learning models also suggest that strategic sampling does diminish monotonically over time. However, even if overall strategic sampling decreases over time, the motivation for sampling for information purposes may increase *periodically* due to the introduction of new generations of the products or the changing needs of the consumers, even if overall uncertainty diminishes. In high-tech product categories, newer generations of the products (e.g., software updates) are very common whereas in frequently-purchased packaged consumers product categories, brands may be re-launched. Of particular interest to us are the changing consumer needs that may prompt increased strategic sampling periodically. In the disposable diapers category, for example, consumers need to switch diaper sizes as their babies grow older.[2] What is common among these examples is that when consumers purchase a specific new size, new generation of a product or the newly relaunched brand, they could expect that the quality is highly correlated with other sizes of the same brand, the older generation product or pre-relaunch brand but not necessarily perfectly so. Under these circumstances, the need to switch to new sizes/generations/relaunched versions of brands may trigger periodic strategic sampling.

The goal of our research is to model how consumers make decisions in product categories where 1) consumers are new to the market and they have little information about alternatives and overall uncertainty is large; 2) there are periodically changing needs (e.g., the need to switch to a larger size when the baby grows out of a size in the disposable diapers category); 3) consumer price sensitivities may change over time as overall uncertainty diminishes.

To accomplish this goal, we propose and estimate a dynamic structural model of demand using scanner data for the disposable diaper category where one can observe the behavior of new consumers (first time parents) over time. The model allows use experience with a brand's

[2] In our research context, size means the physical dimension of one unit of a product (for example, Size 1 (smallest) to Size 5 (largest) of a piece of diaper. Package size, on the other hand, refers to the number of a product of the same size in a package (for example, 16 pieces of Size 4 diaper).

particular size to provide noisy information about another size of the same brand as well. We also allow for changes in consumer price sensitivities when babies grow older and parents know more about the product category. Although our modeling approach would be applicable to any frequently purchased product category where consumer needs change periodically or new product introductions and relaunches are frequent, disposable diapers is an ideal category for our investigation since it provides an ideal context for learning in general and changing needs and price sensitivities over time in particular. First, in the diapers category, potential uncertainty about quality is high because many parents are new to the market and may switch brands to learn about them. Second, unlike in categories such as coffee or detergent where consumers may be using the category for many years and have very well-established tastes, in the disposable diapers category consumer price sensitivities may evolve as well (e.g., consumers may become more price sensitive over time). Thus, by focusing on first-time parents,[3] one can observe the evolution of consumer preferences and choices as the parents get more experienced. Third, as we have mentioned, in the diapers category, the need for strategic sampling may increase periodically due to the need to switch to a bigger size, which is the most distinguishing feature of this category among many other frequently purchased packaged consumer goods.

The forward-looking structural demand model proposed and estimated accounts for changing needs and price sensitivities over time and the possibility that motivation for strategic sampling may increase periodically as the needs change. To our best knowledge, this is the first dynamic structural demand model with forward-looking consumers with experience spill-over effects (that is, experience with a particular size of a brand contains (noisy) information about the other sizes of the brand), and changing price sensitivities. Our results show that 1) consumer experience of a particular size of a particular brand serves not only as a quality signal for that

[3] Focusing on first –time parents also alleviates greatly the initial conditions problem that all dynamic models are subject to.

size but also for other sizes of the same brand; 2) consumer brand-size preference is duration dependent in such a way that it first increases and then decreases with the time that consumers stay with a particular brand-size; and 3) consumer price sensitivities change when their babies grow older. Finally, we conduct policy experiments to describe how marketers may tailor their marketing activities when consumer needs change periodically.

II. The Model

II.1. Overview

We model household behavior in a market in which households may be uncertain about product quality and risk-averse. We allow households to use their use experience as a signal of product quality; that is, they learn about product quality through use experience and update their expectations in a Bayesian manner. However, we go beyond the usual Bayesian learning models in that we account for spill-over effects of the signals from one subset of the product category to another by allowing correlations between signals from different subsets of the category.[4] In our particular study of purchases in the diapers market, we accomplish this by projecting the information provided by use experience in a size of a diaper brand onto other sizes of the same brand. We allow households to be forward-looking in the sense that they maximize the expected value of the aggregate present values of their future utilities over a planning horizon. This leads to strategic-sampling as in Erdem and Keane (1996). However, we go beyond the now standard strategic-sampling model of Erdem and Keane (1996) in that we account for the fact that in certain markets consumer needs may change periodically. In the specific case of diapers category, for example, there is the need to switch to a bigger size as the baby grows older, which

[4] Previous papers that incorporated spill-over learning effects across products or products attributes assumed myopic agents (e.g., Erdem 1998). One exception is Dickenson (2011), who, like us, allows for forward-looking behavior. Dickstein considers a model with forward-looking physicians facing a multi-armed bandit problem, where a physician is uncertain about his patients' intrinsic preference for drugs' characteristics, and he makes use of patients' total utility of consuming a drug in time *t* to update his belief about their preferences. The proposed model does not allow for risk-averse behavior or evolving needs or price sensitivities over time.

may prompt temporary increased strategic sampling. This is so because use experience signals (quality information through use experience) associated with a smaller size of a brand may not be perfectly correlated with use experience signals (quality information through use experience) associated with a larger size. We also allow for price sensitivities to be different for smaller versus larger sizes to capture the possibility that parents' price sensitivities diminish as they get more experienced. A formal description of what we propose is below.

II.2. Consumer Utility

Consider a set of households $\mathrm{H}=\{h \mid h=1,...,H\}$ purchasing from a product category. Let $\mathrm{T}=\{t \mid t=1,...,T\}$ be the time period (week), $\mathrm{J}=\{j \mid j=1,...,J\}$ be the set of brands available in the category, and $\mathrm{K}=\{k \mid k=1,...K\}$ be the set of sizes of all of the brands.[5] The choice set for each household is specified such that household h's choice set does not include all the K sizes since sizes too large or too small would not meet the babies' needs. Specifically, at each *t* when/if a purchase occurs, we assume that the choice set includes the purchased size[6], as well as the adjacent one size for sizes 1 and 5, and the adjacent two sizes for sizes 2, 3 and 4 (that is, a household could consider both size 1 and 2 when they are purchasing size 1; while they could consider size 1, 2, and 3 when they are purchasing size 2; and they would consider size 4 and 5 when purchasing size 5). This choice set of household *h* at time *t* can be written as $\mathrm{K}_{ht}=\{k-1,k,k+1\}\cap\mathrm{K}$, if *k* is the size purchased by the household *h* on purchase occasion *t*. Lastly, the set I_{ht} is the information available to household *h* at time $t \geq 0$. Let us start with utility functions of the form[7]

---

[5] Here size refers to the size of the individual diaper, such as the newborn size (there are 5 total sizes), and not to the package size.

[6] When a household does not make a purchase during week *t*, we assume his choice set included the <u>*last*</u> purchased size and its adjacent sizes as discussed in the above text.

[7] Please note that we tried data on features as a control variable (we do not have display variable). However, the feature variable was statistically insignificant in the structural models we estimated so we did not use it as a control.

$$U_{htjk} = \beta_{hm_k} p_{htjk} + w_h \left(Q_{Ehtjk} - r_h Q^2_{Ehtjk} \right) + \mu_h \left(D_{h,t-1,k} - s_k D^2_{h,t-1,k} \right) + \varepsilon_{htjk} \qquad (1)$$

where brand $j=1,\ldots,J$, size k=(1,2) or (1,2,3) or (2,3,4) or (3,4,5) or (4,5), household $h=1,\ldots,H$, and $t=1,\ldots T$. In the above, p_{htjk} is the price faced whereas Q_{Ehtjk} is the quality experienced by household h of size k of brand j at time t. The parameters w_h and r_h are the respective measures of the utility weight of and degree of risk aversion to the unobservable product quality, both expected to be positive. The parameter β_{hm_k} is the price coefficient, expected to be negative. Here the subscript m_k captures diaper size. We allow for size specific price coefficients and allow these coefficients to be different between small sizes and large sizes. More specifically, m_k will denote a small size when k=1, 2 or 3, say, m_k =1; and m_k will denote a large size when k=4 or 5, say, m_k =2. Thus, rather than allowing a separate coefficient for each size, we lumped them into two groups for parsimony. Allowing for price sensitivities to be size-specific enables us to capture whether and how these sensitivities evolve over time as babies grow (and as parents accumulate more quality information and face reduced uncertainty).[8] ε_{htjk} is the taste shock that becomes known to the household at time t but is unknown to the econometrician. We specify the distributional properties of ε_{htjk} later in the text.

$\mu_h \left(D_{h,t-1,k} - s_k D^2_{h,t-1,k} \right)$ captures duration dependence in a size. D_{htk} is the number of periods in the k^{th} size until time t since the household h made the first purchase of the k^{th} size. We call D_{htk} the size duration variable. This term captures changing needs of consumers over time. In the specific case of diap

ers, we expect that the time spent in a particular size would increase the expected utility initially but that as the baby gets closer to growing out of the current size, the positive influence

[8] Parents may also accumulate more experience with using diapers in general as their kids get older and the babies may become less sensitive to the type of diaper they wear. However, regardless of the multitude of reasons that may lead to differential sensitivities for price over time, allowing for such changes in these sensitivities are important in markets where preferences evolve over time.

of the time spent in the current size would diminish gradually[9] so this specification allows for this possibility. While the duration weight μ_h is heterogeneous across households, the size specific location coefficients s_k are assumed to be the same for all of the households. Provided that the duration weight μ_h is positive in the mean and that $0<s_k<1$, consumer utility increases with the size duration for as long as $D_{htk} <1/ s_k$ and, thereafter, it starts to decrease.[10]

II.2. Consumer Expected Utility

Let Q_{jk} denote the unobserved *true* quality of size *k* of brand *j* about which the *experienced* quality Q_{Ehtjk} fluctuates. Fluctuations of the experienced quality Q_{Ehtjk} about the unobserved true quality Q_{jk} may occur for many reasons. One possibility is the variability of product quality across batches of products, to name one. Another possibility is the context dependence of consumer use experiences, a more plausible explanation for product categories covered by scanner data typically, to name another. And there may be many other reasons. Irrespective of the reasons, however, we formulate the fluctuations in use experience by assuming that each use experience provides a noisy but unbiased signal of quality according to $Q_{Ehtjk}=Q_{jk}+\xi_{htjk}$, here $\xi_{htjk} \sim N(0,\sigma_{\xi}^{2})$ and σ_{ξ}^{2} is the *experience variability* (the reciprocal of which denotes the precision of the use experience information or the quality signal associated

[9] When babies first grow out of a size, the fit to the new size may be not perfect, so utility may first increase as time passes and the fit gets better. Then, when the baby is about to grow out of a size, the fit may diminish.

[10] We also tried a utility specification where Equation 1 has a last brand purchase dummy to capture any one-lag state dependence effects not related to learning (e.g., switching costs (Dube, Hitsch, and Rossi 2009) or preference inertia (Shin, Misra, and Horsky 2010)). Indeed , Osborne (2011) found that in frequently purchased product categories there are both learning and switching costs. The coefficient of the last purchase dummy is identified in such contexts as discussed on Osbourne (2011). We found that lagged purchase dummy has statistically significant but (size-wise) very small effect, and the results were very similar between the two models. We turned that component off for three reasons. First, we turned it off for parsimony since our model has already quite a few "moving-parts" as it focuses on evolution of needs (the need to switch to a different size), learning across sizes and changing price sensitivities. Second, a learning model fits our data better than a model with no learning but a lagged dependent variable or with a weighted average of past purchases variable. Third and most importantly, these lagged purchase variables added to learning models are behaviorally difficult to interpret.

with the use experience). We assume further that ξ_{htjk} are identically and independently distributed.

It is evident that the quality *experienced* by household *h* of size *k* of brand *j* at time *t*, that is, Q_{Ehtjk}, does not need be the same as the quality *perceived* by the same household of the same size of the same brand at the same time, that is, Q_{htjk}. We assume that the *perceived* quality is given by $Q_{htjk}=Q_{jk}+\upsilon_{htjk}$, where the perception errors are distributed as $\upsilon_{htjk} \sim N(0, \sigma^2_{\upsilon_{htjk}})$. The perception variance (variance of quality beliefs) in period *t*, $\sigma^2_{\upsilon_{htjk}}$, is updated after a purchase is made in period $t-1$.

We also allow initial variance of quality beliefs $\sigma^2_{\upsilon_{h0k}}$ to be a function of purchases in the category during the initialization period (in our sample, 27 weeks) as a crude measure of heterogeneity in initial uncertainty across households. Thus, we assume that

$$\sigma^2_{\upsilon_{h0k}} = \lambda_0 + \lambda_1 \left(\text{Number of Purchases in the Initialization Period for Household } h\right)$$

Finally, since we have assumed that the use experience signals are unbiased, we have $Q_{htjk}=E[Q_{Ehtjk}|I_{h,t-1}]= E[Q_{jk}|I_{h,t-1}]$.

Prior to making a purchase decision in period *t*, household *h* forms the $I_{h,t-1}$ -conditional expectations of U_{htjk} for each of the sizes of all brands as follows:

$$\begin{aligned} U_{htjk} &= \beta_{hm_k} p_{htjk} + w_h Q_{Ehtjk} - w_h r_h Q^2_{Ehtjk} - w_h r_h \left(\sigma^2_{\upsilon_{htjk}} + \sigma^2_{\xi}\right) \\ &\quad + \mu_h \left(D_{h,t-1,k} - s_k D^2_{h,t-1,k}\right) + \varepsilon_{htjk} \end{aligned} \tag{2}$$

Since $Q_{htjk}=Q_{jk}+\upsilon_{htjk}$, the $I_{h,t-1}$ -conditional expected utilities $\text{E}[U_{htjk} \,|\, I_{h,t-1}]$ depend not only on the unobservable product qualities Q_{jk} but also on the perception errors υ_{hjkt}. Rewritten explicitly:

$$U_{htjk} = \beta_{hm_k} p_{htjk} + w_h \left(Q_{jk} + \upsilon_{htjk} \right) - w_h r_h \left(Q_{jk} + \upsilon_{htjk} \right)^2 - w_h r_h \left(\sigma^2_{\upsilon_{htjk}} + \sigma^2_{\xi} \right) + \mu_h \left(D_{h,t-1,k} - s_k D^2_{h,t-1,k} \right) + \varepsilon_{htjk} \qquad (3)$$

As is evident from equation (3), there are two sources of consumer uncertainty: the first is the perception variability $\sigma^2_{\upsilon_{htjk}}$ whereas the second is the experience variability σ^2_{ξ}. Although the perception variability diminishes with use experience in our model, the experience variability does not.

Additionally, we specify the utility of no purchase as $U_{ht00} = \gamma_0 + \gamma_h INV_{ht} + \varepsilon_{ht00}$, where INV_{ht} is the household's inventory of the product category. We model inventory as $INV_{ht} = INV_{ht-1} + q_{ht-1} - C_h$, where q_{ht-1} denotes the quantity of category purchased by household *h* at purchase date *t-1* and C_h denotes household *h*'s consumption rate (Bucklin and Gupta 1992). We measure the consumption rate, C_h, as the average weekly consumption of diaper and it is computed as the total number of pieces of diaper purchased by household h divided by the number of weeks in the sample period. Lastly, γ_h is the household-specific inventory weight and γ_0 is the intercept of the no-purchase utility.[11] [12]

Finally, we let β_h, w_h, and r_h be heterogeneous across consumers and, following Heckman (1981), adopt a latent class approach. (β_m, w_m, r_m) as well as the associated population type proportions π_m for each of the consumer segments *m=1,2,…,M*. Many papers that involve Dynamic Programming models utilize a latent class approach to capture unobserved

[11] The estimate of γ_0 was not statistically significant so we did not report them in the result tables.

[12] We should also note that we model brand choice and purchase incidence but do not model quantity choice (rather we model the impact of inventories on the probability of purchase incidence in a reduced-form way). Previous papers on forward-looking dynamic structural models focused either on quality expectations, learning and strategic sampling in the context of brand choice or on both brand and quantity choice and price expectations but assumed away quality learning and strategic sampling since it is not feasible to model both processes in one model that explicitly allows for both quality and price expectations (Erdem, Keane and Sun (2008)). Furthermore, Ching, Erdem and Keane (2012) do so in a semi-structural model and find that in the diapers category the quality learning effects are significant whereas the price expectation effects are not.

heterogeneity since imposing a continuous distribution for heterogeneity would imply solving the Dynamic Programming problem for each household which is difficult to accomplish.[13]

II. 3. Consumer Learning

We assume for each household $h \in \mathrm{H}$, brand $j \in \mathrm{J}$ and size $k \in \mathrm{K}$ that the initial perception errors υ_{h0jk} are correlated across sizes and their correlation matrix is given by:

$$R = \begin{bmatrix} 1 & \rho_{12} & 0 & . & . & . & . & 0 \\ . & 1 & \rho_{23} & . & . & . & . & 0 \\ . & . & 1 & \rho_{34} & . & . & . & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & 1 & \rho_{K-2,K-1} & 0 \\ . & . & . & . & . & . & 1 & \rho_{K-1,K} \\ . & . & . & . & . & . & . & 1 \end{bmatrix}$$

where $\rho_{h0kl} = \rho_{kl}$ are the initial correlations between sizes *k* and *l*. This specification indicates that only the adjacent sizes are correlated and that the initial size correlations are uniform across brands and households, by assumption. We denote by ρ_{htkl} the time $t \geq 1$ correlation coefficients between the sizes of the brands, which are updated over time as we describe below.

Since υ_{htjk} are correlated across sizes, we have the following relationships across different sizes of the brands for any $t \geq 1$:

$$Q_{Ehtjk} = \kappa_{htjk|l} Q_{Ehtjl} + \eta_{htjk|l} 1_{\{l \neq k\}}, \quad \eta_{htjk|l} \sim N\left(0, \sigma^2_{\eta_{htjk|l}}\right) \tag{4}$$

In mathematical terms, this is the linear projection of one vector on another, and κ can be thought of the ordinary least square coefficient for Q_{Ehtjl}. Therefore,[14]

$$\kappa_{htjk|l} = \frac{\sigma_{\upsilon_{h,t-1,jk}} \sigma_{\upsilon_{h,t-1,jl}} \rho_{h,t-1,kl}}{\sigma^2_{\upsilon_{h,t-1,jl}} + \sigma^2_{\xi} 1_{\{l \neq k\}}} \tag{5}$$

[13] Hartmann (2006) and (2010) are two exceptions that allow for richer unobserved heterogeneity structures.
[14] More details of the learning model are provided in a Technical Appendix available upon request from the authors.

and

$$\sigma^2_{\eta_{htjk|l}} = \frac{\left(\sigma^2_{\upsilon_{h,t-1,jk}} + \sigma^2_{\xi} 1_{\{l\neq k\}}\right)\left(\sigma^2_{\upsilon_{h,t-1,jl}} + \sigma^2_{\xi} 1_{\{l\neq k\}}\right) - \sigma^2_{\upsilon_{h,t-1,jk}} \sigma^2_{\upsilon_{h,t-1,jl}} \rho^2_{h,t-1,kl}}{\sigma^2_{\upsilon_{h,t-1,jl}} + \sigma_{\xi}{}^2 1_{\{l\neq k\}}} \tag{6}$$

whereas $1_{\{A\}}$ is the indicator function that returns 1 if the statement *A* is true and 0 otherwise.

It is clear from equations (3) and (4) that as long as ρ_{htkl} are non-zero, use-experience in a particular size of a brand provides information for other sizes of the same brand as well. Our approach could be employed to study the spill-over effects of signals from one domain to another in the context of forward-looking consumers. As previously indicated, all previous learning models with spill-over effects have assumed myopic consumers, with the exception of a working paper by Dickenson (2011), whose method requires risk-neutrality (and Dickenson's model does not allow for changing needs and price sensitivities). Our model allows for risk-aversion as well, which has been shown to hold empirically in previous learning papers. Furthermore, it is clear from equation (4) also that $0 < \left|\kappa_{htjk|l}\right| < 1$ when $k \neq l$ whereas $\kappa_{htjk|k} = 1$. This implies that the noisy information provided by the use experience in a different size of a brand is less than the information provided by the use experience in the current size.

Let us now denote by J_{ht} and K_{ht} the respective brand and size purchases of household *h* on purchase occasion *t*. With this definition and the above relationships between the sizes *k* and *l*, we are now ready to describe how the consumers update their perceived qualities in our model. We assume that after buying size *l* of brand *j* on purchase occasion *t*, household *h* updates the priors about the mean quality of size *k* of brand *j* using Bayesian updating rules (see, for example, DeGroot 1970). Therefore, we have the following updating equations for $t > 1$:

$$\upsilon_{htjk} = \kappa_{htjk|K_{h,t-1}} \upsilon_{h,t-1,jk} + \frac{\kappa^2_{htjk|K_{h,t-1}} \sigma^2_{\upsilon_{h,t-1,jk}}}{\kappa^2_{htjk|K_{h,t-1}} \left(\sigma^2_{\upsilon_{h,t-1,jk}} 1_{\{J_{h,t-1} \neq j\}} + \sigma^2_\xi \right) + \sigma^2_{\eta_{htjk|K_{h,t-1}}} 1_{\{K_{h,t-1} \neq k\}}} \cdot \left\{ \kappa_{htjk|K_{h,t-1}} \left(\xi_{htjk} - \upsilon_{h,t-1,jk} \right) + \eta_{h,t-1,jk|K_{h,t-1}} 1_{\{K_{h,t-1} \neq k\}} \right\} 1_{\{J_{h,t-1} \neq j\}} \quad (7)$$

$$\sigma^2_{\upsilon_{htjk}} = \frac{\kappa^2_{htjk|K_{h,t-1}} \left(\kappa^2_{htjk|K_{h,t-1}} \sigma^2_\xi + \sigma^2_{\eta_{htjk|K_{h,t-1}}} 1_{\{K_{h,t-1} \neq k\}} \right) \sigma^2_{\upsilon_{h,t-1,jk}}}{\kappa^2_{htjk|K_{h,t-1}} \left(\sigma^2_{\upsilon_{h,t-1,jk}} 1_{\{J_{h,t-1} \neq j\}} + \sigma^2_\xi \right) + \sigma^2_{\eta_{htjk|K_{h,t-1}}} 1_{\{K_{h,t-1} \neq k\}}} \quad (3)$$

and

$$\rho_{hjklt} = \begin{cases} \rho_{hjkl,t-1} & J_t \neq j \\ \dfrac{\mathrm{cov}\left(\upsilon_{hjkt}, \upsilon_{hjlt}\right)}{\sigma_{\upsilon_{hjkt}} \sigma_{\upsilon_{hjlt}}} & J_t = j \end{cases} \quad (4)$$

where,

$$\mathrm{cov}\left(\upsilon_{hjkt}, \upsilon_{hjlt}\right) = \left(1 - \beta_{htjk|K_{h,t-1}}\right)\left(1 - \beta_{htjl|K_{h,t-1}}\right) \kappa_{htjk|K_{h,t-1}} \kappa_{htjl|K_{h,t-1}} \sigma^2_{\upsilon_{h,t-1,jk}} + \beta_{htjk|K_{h,t-1}} \beta_{htjl|K_{h,t-1}} \kappa_{htjk|K_{h,t-1}} \kappa_{htjl|K_{h,t-1}} \sigma^2_\xi, \quad \text{and} \quad (5)$$

$$\beta_{htjk|K_{h,t-1}} = \frac{\kappa^2_{htjk|K_{h,t-1}} \sigma^2_{\upsilon_{h,t-1,jk}}}{\kappa^2_{htjk|K_{h,t-1}} \left(\sigma^2_{\upsilon_{h,t-1,jk}} 1_{\{J_{h,t-1} \neq j\}} + \sigma^2_\xi \right) + \sigma^2_{\eta_{htjk|K_{h,t-1}}} 1_{\{K_{h,t-1} \neq k\}}}$$

We observe from the above that both the above variances and correlations are decreasing to zero with the number of purchases, as is evident from equations (5) through (9). Of course, this does not mean that consumers will eventually learn the product qualities of the brands and sizes with certainty since that would require infinitely many purchases of diapers from each brand and size.

II.4. Consumer Expected Utility Maximization over the Planning Horizon

Recalling that J_{ht} and K_{ht} are the respective brand and size purchases of household *h* on purchase occasion *t*, we suppose that the forward-looking household *h* solves the following dynamic programming problem:

$$\max_{\{(J_{ht}, K_{ht}) \in J \times K \cup (0,0) | t=1,2,\ldots,T_H\}} E\left[\sum_{t=1}^{T_H} \beta^{t-1} E\left[U_{htJ_{ht}K_{ht}} \mid I_{h,t-1} \right] \mid I_{h0} \right] \quad (6)$$

where β is the one-period discount factor. We choose β=0.995.[15] The planning horizon T_H may go beyond the end of observation period T; that is, we may have $T_H > T$.

To solve the consumer dynamic programming problem (11), we apply Bellman's principle to solve this problem by finding value functions corresponding to each alternative choice. The value of choosing alternative *jk* at purchase occasion *t* is:

$$V_{htjk}\left(I_{ht}\right) = E\left[U_{htjk} \mid I_{ht}\right] + \beta \cdot E\left[\max_{lm} V_{h,t+1,lm}\left(I_{h,t+1}\right) \mid I_{ht}\right] \quad (7)$$
$$\text{where } j = 0,...,J,\ k = 1,...,5$$

subject to the terminal condition

$$V_{hT_H jk} = \mathrm{E}\left[U_{hT_H jk} \middle| I_{h,T_H-1}\right] \quad (8)$$

To compute the V_{htjk} , we must compute the above Emax functions appearing in recursion relation (Equation (12)) for each of the alternatives. This is not an easy task. However, if we assume that the stochastic taste shocks ε_{hjkt} to the expected utilities $E[U_{hjkt} \mid I_{ht}]$ are identically and independently extreme value distributed, then we can obtain closed form expressions for the above $\left\{I_{h,t-1}, J_{ht} = j, K_{ht} = k\right\}$ conditional Emax functions as detailed in Rust (1987). We assume now that this is the case.

II. 5. Consumer Choice Probabilities and the Likelihood Function

Since we have assumed that the stochastic taste shocks ε_{hjkt} are identically and independently extreme value distributed, the consumer choice probabilities are the conditional logit probabilities of McFadden (1974). When consumers are myopic, the period *t* choice

---

[15] We fixed the weekly discount factor β at 0.995 since the discount factor is often difficult to identify even when certain variables can be found that affect expected payoffs but nor current utility (that is, exclusion restrictions may exist). For example, Erdem and Keane (1996) found, in a similar but simpler model, the likelihood was quite flat over a range of discount factors in the vicinity of 0.995, which was the case for us too. We estimated the model with weekly discount factors in the vicinity of 0.999 but the results were not very sensitive to the exact value of the discount factor. Please note that the best way to identify the discount factor is either to find contexts where proper exclusion restrictions and practical identification exist (e.g., Chung, Steenburg, and Sudhir (2013)) or use additional (experimental or field) data to pin down the discount factor (e.g., Yao, Mela, Chiang, and Chen (2012)) but we do not have such data. There are indeed very few cases where such data are available.

probabilities of household *h* for latent class *m* conditioned on the period *t* perception errors vector $\upsilon_{ht} = \left\{ \upsilon_{htjk_{ht}} | \left(j,k \right) \in \mathrm{J} \times \mathrm{K}_{ht} \cup \left(0,0 \right) \right\}$ (where $\mathrm{K}_{ht} = \{k-1, k, k+1\} \cap \mathrm{K}$) are:

$$\theta_{htjk}\left(\Theta_m | \upsilon_{ht}\right) = \frac{\exp\left\{E\left[U_{htjk} | I_{h,t-1}\right]\right\}}{\sum_{(m,n)\in J\in K_h \cup (0,0)} \exp\left\{E\left[U_{htmn} | I_{h,t-1}\right]\right\}} \tag{9}$$

while when consumers are forward-looking, they are

$$\theta_{htjk}\left(\Theta_m | \upsilon_{ht}\right) = \frac{\exp\left\{E\left[U_{htjk} + \beta \mathrm{E}\left[V_{h,t+1} | I_{h,t-1}, J_{ht} = j, K_{ht} = k\right] | I_{h,t-1}\right]\right\}}{\sum_{(m,n)\in J\in K_h \cup (0,0)} \exp\left\{E\left[U_{htmn} + \beta \mathrm{E}\left[V_{h,t+1} | I_{h,t-1}, J_{ht} = m, K_{ht} = n\right] | I_{h,t-1}\right]\right\}} \tag{10}$$

where $\Theta_m = \left\{ \beta_m, \phi_m, r_m, w_m, \sigma_\xi, \sigma_\upsilon, \sigma_\eta, \varphi, Q, \rho, s \right\}$ is the class *m* parameter vector in which $Q = \left\{ Q_{jk} \left| (j,k) \in \mathrm{J} \times \mathrm{K} \right. \right\}$, $\rho = \left\{ \rho_{kl} \left| l = k+1; k = 1, K-1 \right. \right\}$ and $s = \left\{ s_k \left| k = 1, K \right. \right\}$.

Irrespective of whether households are myopic or forward-looking, however, the class *m* $\upsilon_h = \left\{ \upsilon_{ht} \left| t \in \mathrm{T} \right. \right\}$ conditional likelihood function of household *h* associated with the purchases made over the observation period T is:

$$L_{hm}\left(\Theta_m | \upsilon_{ht}\right) = \prod_{t=1}^{T} \prod_{(j,k)\in J\times K \cup (0,0)} \theta_{htjk}\left(\Theta_m | \upsilon_{ht}\right)^{Y_{htjk}} \tag{11}$$

where $Y_{htjk} = 1$, if household *h* bought size *k* of brand *j* on purchase occasion *t*, while $Y_{htjk} = 0$ otherwise.

Had the consumer perceptions errors vector υ_h been observable to the econometricians, the above υ_h conditional likelihood function of household *h* for each of the latent classes *m =1,2,…,M* would have sufficed. However, because the consumer perceptions errors vector υ_h is not observable to us, we need to work with the unconditional likelihood function for latent class *m* given by:

$$L_{hm}\left(\Theta_m\right) = \int_{\upsilon_h \in \Omega} L_h\left(\Theta_m \left| \upsilon_h\right.\right) f(\upsilon_h) d\upsilon_{h1} d\upsilon_{h2} ... d\upsilon_{hN} \tag{12}$$

where $f(\upsilon_h)$ is the joint distribution, N is the length, and Ω is the obvious domain of the household perception errors vector υ_h. Since it is impossible to carry out the above integration analytically, we have to resort to numerical techniques. We use the interpolative regression method developed by Keane and Wolpin (1994) and used in many previous forward-looking learning models, including Erdem and Keane (1996).

Once $L_{hm}\left(\Theta_m\right)$ is computed by employing the simulation technique mentioned above, we can then calculate the likelihood function of household *h* from:

$$L_h\left(\Theta\right) = \sum_{m=1}^{M} \pi_m L_{hm}\left(\Theta_m\right) \tag{13}$$

where $\Theta = \left\{\left(\Theta_m, \pi_m\right) \middle| m = 1, 2, ..., M\right\}$ is the overall parameter vector. The likelihood function of the entire sample we maximize to estimate the parameter vector Θ is then:

$$L\left(\Theta\right) = \prod_{h=1}^{H} L_h\left(\Theta\right) \tag{14}$$

III. Data and Identification

We use scanner panel data from a large national grocery chain on household purchases of disposable diapers between May 2005 and May 2007 in one store in the San Francisco Bay area. The store is located in a mountainous area and has no other large grocery competitors (stores from other retailers or from the retailer itself) or grocery supercenters (e.g. Target or Wal-Mart) within a 5-mile radius. One potential problem of using data from one retail chain is that the observations of consumers shopping in competing stores are unavailable. Using data from one store that has no competing stores nearby, however, helps to reduce such possible bias. The data record all store visits for 105 weeks in 2005-07 for a panel of over 50,000 households in Northern California. Both the brand purchased and price paid are recorded.

The disposable diapers category is an ideal category for our purposes since: 1) the potential for strategic sampling is high in this category as we are studying the choices of first-time parents who are identified by household demographics information (e.g. children's ages and the number of children) and 2) this is a category where household needs change periodically due to the need to switch to a bigger size when the babies get older, 3) initial conditions problem that is relevant for all brand choice models but even more problematic for learning models is less an issue here since the sample of households analyzed are new to this market as described more in detail below.

We analyze three brands (Pampers, Huggies, and the store brand, which together have a 98% market share) and use the purchase selection procedure (Gupta et al. 1996) to retain households purchasing only these 3 brands in the category. To capture learning behavior over time and minimize initial conditions problem, we then focus on first-time parents who have made at least 22 purchases in 105 weeks. Given that mean purchase frequency is about three to five weeks (it varies by size), this selection criteria allows us to exclude first-time parents whose child is about to grow out of diapers. Among the resulted 1007 regular diaper-user households who made at least 22 purchases, we identify 365 households who are first time parents. We identify first-time parenthood using the total number of children and number of children in each age bracket information available in the data. We define the first-time parents as parents who have 1) *only one* child; and, 2) the child is under 3. Then, we then use a random number generator to assign these households to the calibration and validation samples. In this way, 191 first-time parent households are selected for calibration and 174 for validation.

The observations in the first 27-week (the initialization period) are used to allow for heterogeneity in consumer initial uncertainty. As indicated in Section II.2., we allow initial perceived variance to be a function of purchases in the category during the initialization period. As the estimation sample covers 78 weeks, the calibration and holdout samples have 9102 and 7918 observations, respectively. In the calibration sample, the mean number of diaper purchases is 25 times in a two year period.

Table 1 reports descriptive statistics for different brands and sizes for the calibration sample.

~Table 1 About Here ~

Note that Pampers and Huggies each has 47~48% of the total market share, and the store brand has a 5% market share. Diaper sizes 3 and 4 have the highest market shares (28~29%) while size 1 has the lowest market share at 12%.

Table 2 shows that Pampers is also the highest priced brand with Huggies being a close second and the store brand being the cheapest brand (at a price level that is on average 30% lower than the two national brands).

~Table 2 About Here~

Identification

Table 3 is the switching matrix between consecutive purchases among different sizes of different brands.

~Table 3 About Here~

Table 3 shows that repeat purchase of the same size of the same brands accounts for a large percentage of all purchases, while in the meantime, there are also a significant number of switches between different brands of the same size and switches between the adjacent sizes of different brands. In addition, there are also a significant number of switches across different brands in larger sizes of diapers. These switching patterns and variation across households in regard to purchase histories and switching patterns, as well as price variation over time, aid in identifying the cross-size learning effects, the duration-dependence effects in utility and price sensitivities for smaller versus larger brands. A detailed discussion of how the individual learning parameters are identified in learning models is available in Crawford and Shum (2005), Erdem, Keane and Sun (2008) and Ching, Erdem and Keane (2012).

We should stress here again that the standard learning models (e.g., Erdem and Keane 1996) imply that consumers will settle on small subset of brands once consumers are less uncertain about (a subset) of brands, *ceteris paribus.* Thus, standard leaning models would suggest that we should observe less switching over time as consumers get more experienced. In the diapers case, this would mean, for example, we would observe less switching for larger sizes than smaller sizes. However, if needs change (e.g., need to switch to a bigger size) and this may lead to a temporary increase in switching around the size change time, the smoothly declining switching pattern would not be observed, even if one would still expect overall there would be less switching as time passes by. More importantly, if consumers become more price sensitive when they learn more about the brands in the category, switching due to price promotions may increase, which may more than offset the declining switching due to learning. Indeed, when we calculated switching matrices for each size, we did *not* observe such a declining trend in switching. Instead, when we calculated the percentage of brand switches between brands in Table 3, we find the percentages of brand switching observations to be 11.1%, 10.4%, 15.3%, 19.6%, and 19.0% for Size 1 through Size 5, respectively, indicating an *increasing* percentage of brand switches when size increases.[16]

---

[16] The increased number of switches in larger sizes would occur if the impact of increased price sensitivity dominates the effect of diminished overall strategic sampling on brand switching. To again check data patterns, we categorized the switching observations into two groups: when a household switched to a different brand when the price of the brand switched to is at least 5% lower than its mean price, we categorized the brand switch observations

Finally, to see whether there is any evidence in the data for possible temporary increase in strategic sampling due to size switches, we calculated brand switching conditional on size switching for each size (that is, count all brand switches when a (an adjacent) size switch was made and divide it by the total number of size switches between adjacent sizes irrespective of whether there was brand switching or not). These conditional brand switching percentages (conditional on a size switch) were 61.2%, 68.8%, 59.9%, 53.3%, 48.4%, for sizes 1 through 5, respectively. These are a lot higher than the unconditional brand switches we reported (11.1%, 10.4%, 15.3%, 19.6%, and 19.0% for Size 1 through Size 5, respectively). Thus, data patterns are consistent with the notion that people may be motivated doing strategic sampling when they are switching sizes. We should also stress that 1) the conditional brand switching percentage are going down with the growing size of the diaper, consistent with the idea that there is always more return on trial if one tries early on, 2) overall brand switching is going up, which would be inconsistent with a pure learning model but is consistent with our model that combines learning with changing price sensitivities.[17]

Finally, we list briefly our formal identification restrictions here again: we set $Q_{35}=1$ (i.e., Store Brand Size 5 quality = 1) and measure quality of other brands relative to this product. Furthermore, in the latent-class model, we set the quality weight parameter (w_k) in one of the segments to be 1 as commonly done in this literature (e.g., in a similar vain, Erdem (1998) fixes

as brand switching due to price promotion. Otherwise, we classified the brand switch as a not price promotion related brand switching (which could be due to strategic trial or other reasons). The size-specific brand switching observations categorized as “price promotion related ” yielded percentages of brand switching to be 44%, 48%, 51%, 59% and 63% for Size 1 through Size 5, respectively. That is, while 44% of all size 1 brand switching was “price promotion related” (and 56% was “non-promotion-related), 63% of total size 5 brand switching was “price promotion-related”. Thus, the data patterns suggest non-promotion related brand switches decline over time relative to price promotion related switches. Thus, consumers switch early on more for non-price related reasons (e.g. for strategic trial) while they will switch due to price variation in later periods.

[17] As a final check in regard to the plausibility of our model or consistency with the patterns available in the data, we also obtained reduced form evidence for our duration dependent size utility specification. More specifically, we ran a simple MNL model without learning but with 1) a brand loyalty term as in Guadagni and Little (1983), which represents state-dependence (Che, Sudhir and Seetharaman (2007)) and is a reduced-form way of capturing learning behavior; and, 2) duration in size utility, specified exactly the same way as in equation 1. We find the brand loyalty term is positive, which implies that consumers exhibit positive state-dependence. We also find evidence for the quadratic specification of duration-dependence in our utility function.

the variance of the heterogeneity distribution of the utility weights). Although there is no formal reason for this restriction, it is difficult to estimate the model without this restriction as the likelihood tends to be too flat. Keane (1992) calls this fragile identification since although a parameter may be formally identified, it may be impossible to estimate in practice because the likelihood is too flat; thus, the analysts may need a very (prohibitively) large number of observations in practice.

IV. Empirical Results

IV.1. Model Fit and Model Selection

Our model allows for heterogeneity in the price coefficient (β), duration weight (μ), utility weight on quality (w), risk coefficient (r), and the coefficient for inventory in the no-purchase utility (γ) so we must first choose the number of segments M. We estimated models with 1, 2, and 3 segments and report measures of fit in Table 4.

~Table 4 About Here~

In the best-fitting model (forward-looking consumer choice with across-size learning, duration-dependence in utility, as well as size-specific quality and price parameters), increasing K from one to two improves AIC and BIC by 136 and 87 points, respectively; when we increase the number of segments from two to three, AIC and BIC increase by 80 and 23 points, respectively. While when we increase the number of segments further to four, the BIC did not improve. Based on these results, we decided to use the three-segment model for further analysis.

Comparing the fits of different models we estimated, we can also find out the relative importance of different components of our model. Compared to the model with only learning of brand-size quality (Model 1), modeling consumer learning across adjacent sizes (Model 2) improves the BIC by 164 points while adding duration dependence in the utility specification (Model 3) improves the BIC by 110 points. This shows the importance of accounting for both size learning and duration-dependence in utility in modeling consumer brand choice. The longer a consumer stays with a specific size of a brand, the more likely its utility will decrease. Table 4 also shows that the 3-segment myopic learning model (Model 3) has a log-likelihood that is 393 points worse than the 3-segment model with forward-looking consumers (Model 4), and the 3-segment myopic learning model with size-specific parameters (Model 5) has a log-likelihood that is 52 points worse than the 3-segment model with size-specific parameters and forward-looking consumers (Model 6). Thus, the forward-looking aspect of the model (i.e., strategic trial purchases) is important.

Comparing Model 4 (forward-looking learning model without size-specific price parameters) with Model 6 (forward-looking learning model with size-specific price parameters), we find Model 6 improves the BIC by 39 points. This implies that for the consumers in our diaper purchase data, the price sensitivities change with time. Finally, the best fitting model (Model 6) has adjacent size learning spill-over effects, duration-dependent size utility and forward-looking consumers (as well as price sensitivities that differ for small versus large sizes). As we previously indicated, this model allows for the possibility that consumers may temporarily increase strategic trial around the time they switch sizes.

IV.2. Parameter Estimates

We report parameter estimates for our preferred (three-segment) model (Model 6) in Tables 5(a) and 5(b).

~Tables 5(a), 5(b) About Here~

Table 5 (a) lists the estimates that are homogenous across segments. The quality estimates are all statistically significant.[18] We find the estimates of the four correlation

---

[18] The “true qualities” can be different across segments if there is a baby-diaper match issue. The same issue holds in many other categories as well though and it is not feasible to estimate a dynamic structural model that allows true qualities varying by households. Furthermore, even if the match issue exists, there is no reason to believe that not modeling it would bias the results in a systematic way.

coefficients (ρ) across different sizes to be all positive (between 0.445 and 0.512) and significant. These are the correlation coefficients for the initial perception errors across sizes; therefore, positive estimates indicate that consumers learn through consumption across sizes. Thus, use experience with a specific size gives information about the quality of its two adjacent sizes as well.

The two parameters of the prior quality perception standard deviation (σ_υ) specification are 9.865 for the intercept (λ_0) and 1.862 for the number of purchases in the initialization period coefficient (λ_1). This suggests that quality uncertainty exists and it is a function of number of past purchases in the category. The estimate of the standard deviation of the experience variability (σ_ξ) (capturing the noise in the consumption experience as a source of quality information) is statistically significant and 1.435. Thus, use experience provides (noisy) information about quality. The corresponding variability (σ_η) for across size experience signals is statistically significant as well and is 14.016, suggesting that use experience associated with a brand-size provides noisy information about the adjacent sizes of the same brand. As one would expect, the estimate of (σ_ξ) is lower than that of (σ_η) since quality information about a specific brand size obtained through use experience is expected to be less noisy than quality information obtained about that specific brand size through experience of an adjacent size of the same brand.

Table 5 (b) lists the estimates that differ by segments. Size duration weight parameter μ_h is positive for all three types of consumers. The estimates of the five location parameters s_k are also positive and bound between 0 and 1 (between 0.302 and 0.742), and they are also highly significant except for s_1 (location parameters s_k are homogenous across segments and are given in table 5 (a)). Given the quadratic specification of duration-dependence in our utility function, the above estimates imply that the utility increases with the size duration for as long as $D_{hkt} < 1/s_k$, and thereafter it starts to decrease. Interestingly, the estimates s_k are larger for smaller sizes and become smaller for larger sizes. This implies that consumer utilities decrease with the time spent on a size, more quickly so for smaller diaper sizes than for larger diaper sizes.

The price coefficients are found to be negative and significant for two of the three segments, which account for 97% of the households in our sample, in the full model (Model 6). We find the consumer's price coefficients are higher for the larger size than for the smaller size. In segment 1 (22% of the sample), the price coefficient is -1.428 for the larger size and -0.410 for the smaller size; while in segment 2, the largest segment (76% of the sample), the price coefficient is -8.587 for the larger size and -3.951 for the smaller size. The price coefficient estimates from both segments imply that first-time parents are less price-sensitive when they start buying diapers (small size) for their children, while they become more price-sensitive after they buy the diapers for a while and start buying larger sizes of diapers. Finally, we also find the consumer's quality weight coefficients are positive and significant for all three segments.

It might be useful to compare our results with those reduced-form results in Heilman et al (2000). Different from Heilman et al (2000), we find consumers become more, rather than less, price sensitive when they are buying larger sizes of diapers. In their study, the authors do not have a priori prediction that the price sensitivity will become higher or lower. We think that the finding of increasing price sensitivity when households' babies grow older and parents accumulate more information is intuitively appealing. It also explains better the fact that store brands market shares for larger size diapers are higher than those of smaller size diapers, which is the case both in our and Heilman et. al.'s data.

Overall, different from the reduced form results from Heilman et al (2000) and many papers in the structural dynamic consumer choice literature, the evidence we obtain from the parameter estimates suggests that in the diapers category, consumers do not settle on one brand or a very small set of brands in the diapers category due to learning effects (also labeled as familiarity effect in the reduced-from literature). This is because price sensitivities increase when consumer uncertainty decreases over time. Furthermore, our model implies that strategic sampling may go up temporarily when the household is ready to switch sizes.[19]

V. Policy Experiments

We run a few policy experiments (using the estimates from the preferred model, i.e. Model 6) to explore the implications of our proposed and estimated model.

First, we study κ's impact, (Equations (4) and (5)), which are brand- and size-specific parameters that measure the spillover effect of learning across adjacent sizes when consumer purchase a specific brand-size. In Table 6(a), we sequentially set the values of κ's of different brands to be zero and investigate its effect on the choice probabilities of the brand under consideration and competing brands.

~Table 6(a) About Here~

Since our estimates of size correlation coefficients, ρ's, are positive, the spillover effects measured by κ's are also positive. We find when the value of κ is set to zero, the choice probability for the brand under consideration goes down while the competing brands' choice probabilities go up. This is consistent with our finding of positive spillover effects of brand-size learning across adjacent sizes of the same brand. We also find the decreases in choice probabilities are much larger for Pampers and Huggies than for the store brand, indicating a much higher positive spillover effect for the two national brands.

In Table 6(b), we investigate the effects of correlation between initial uncertainties on choice probability. Estimates of correlation coefficients, ρ's, between initial perception errors across sizes are positive. In this analysis, we first calculate the base choice probabilities of different brand-sizes under the scenario that all the correlation coefficients, ρ's, are zero. Then

[19] This implication is consistent with data that show that there is indeed more brand switching conditional on size switching.

we sequentially turn on the ρ 's using our estimates and recalculate the choice probabilities. We find, with few exceptions, that an increase of size correlation coefficient from zero to positive values leads to higher choice probabilities for both adjacent sizes.

~Table 6(b) About Here~

Next, in Table 7, we simulate a 10% price cut for the three brands that last for 1 week, and then calculate the cumulative change in consumer's choice probability over that week and the next 9 weeks. To distinguish the difference in consumer responses to price promotion of small and large sizes, we do two simulation exercises, one each size, and compare the results.

~Table 7 About Here~

We have two important findings: 1) Consumer's own price elasticity for the large size is higher than that for the smaller size. For example, after a 10% price cut, the own price elasticity for the large size of the store brand (the economy brand) is roughly 4.3, while it is 3.4 for the small size; similar results are found for the premium brands: Pampers and Huggies; 2) Consumer's cross price elasticity for the large size is higher than that for the smaller size. Here, the most interesting finding is for the store brand. We find, after a 10% cut of the store brand's price, the cross elasticities for Pampers and Huggies with respect to the store brand are roughly -1.24 and -1.27 for the large size, while they are roughly -0.41 and -0.50 for the small size. Interestingly, the cross price elasticities for the store brand (economy brand) with respect to Pampers and Huggies are also larger for the large size, but the magnitude of increase is much smaller than those for Pampers and Huggies with respect to the store brand. In other words, when the store brand is offering a price cut of the large size of diapers, consumers are more likely to switch from premium (high quality) brands to economy (low quality) brands than they were when buying smaller brands.

Lastly, we conduct policy experiments that examine the impact free samples have on consumer choices. The free sample implies that a consumer gets the chance to try a brand for free, and the simulations conducted assume that all the consumers get the free sample and everybody uses (consumes) it. These free sample policy experiments can further shed light on the cross-size learning effects on consumer choice. In addition, we can also investigate whether providing free samples will lead to more strategic sampling from consumers on a particular brand as compared to other brands.

Tables 8(a) and 8(b) report the actual frequency of purchases for the 15 brand-sizes and baseline simulation results for 25 purchase occasions.

~Tables 8(a) and 8(b) About Here~

The *i*-th row of the table under the heading "Purchase Frequency" reports the number of households that bought a given brand-size during their *i*-th purchase occasion. For example, store brand size 1 was purchased by 1 household on its first purchase occasion, by 1 household on its second purchase occasion, and so on. Baseline simulation results are obtained from the proposed model (Model 6) using the estimated parameters. As can be seen from these two tables, baseline simulation results approximate the observed purchase frequencies well, especially for frequently bought brand-sizes.

To assess the impact of alternative strategy changes, the baseline simulation results (Table 8(a) and Table 8(b)) must be compared with post-intervention (the distribution of the free sample) figures. In Table 8(c), we report the post-intervention purchase frequencies after providing free samples of the small or large size of different brands of diapers to households at the end of the first week. For comparison purposes, we calculate these frequency values from 1) our proposed full model (Model 6 as in Table 4); 2) a nested model where we do not have size-specific price coefficients (Model 4 as in Table 4). 3) a nested model where the duration-dependent in utility component is turned-off (Model 6a); and finally, 4) a nested model where both the duration-dependent in utility component and spill-over effects are turned-off (Model 6b).

~Table 8(c) About Here~

The results in Table 8(c) indicate that for our proposed model, free samples of the small size (sizes 1~3) provided by a brand before the second purchase occasion increase sales of that brand for sizes 1~3 and their adjacent size: size 4. Similarly, free samples of the large size (sizes 4~5) provided by a brand also increase sales of that brand for sizes 4~5 and their adjacent size: size 3. This could be due to the existence of both spillover effect (across sizes) and duration-dependent utilities.

First we compare the results from the proposed full model (Model 6 as in Table 4) to Model 4 (as in Table 4, which has no size-specific price parameters). When there are no size-specific coefficients as in Model 6, the price coefficient estimates from Model 4 are larger in magnitude for the small size (and *vice versa* for the large size). The simulation results show that the 1) for the small size, the increases in sales are *larger* for the brands which provide free samples in Model 4 than in Model 6; 2) for the large size, the increases in sales are *smaller* for the brands which provide free samples in Model 4 than in Model 6. These two findings are consistent with the implications from parameter estimates, and confirm the importance of accounting for size-specific price coefficients.

Next, we compare the results from the proposed full model (Model 6 as in Table 4) to those from the other two nested versions of the model (forward-looking model without size duration-dependent in utility (Model 6a) and forward-looking model without duration-dependent in utility and across-size learning spillover effects (Model 6b)). Compared to the results from the proposed model, we find that 1) in Model 6a, the increase in sales are more for the sizes for which free samples were provided than for the adjacent sizes and other sizes, especially for the large size; and 2) in Model 6b, the increase in sales are only for the sizes for which free samples were provided and there is no increase in sales for the adjacent sizes. This implies that without size-learning spillover and duration-dependent in utility effects, free samples will only benefit the sizes in which free samples are given. When there are only spillover effects but no duration-dependence in the utility function, the magnitude of the increase in sales will be smaller for the adjacent sizes (or none for the other sizes) since households will not expedite their switching to the next available size as in the case of duration-dependent utility.

VI. Discussion and Conclusions

We estimated for the first time in literature a forward-looking structural dynamic learning model where the need for learning and strategic sampling may increase periodically, there are spill-over learning effects across the sizes of a brand and consumer and price sensitivities may not be fixed over time (that is, they may differ for small versus large size diapers). Our estimation results show: 1) consumer experience of a particular size of a particular brand serves as a quality signal of the other (adjacent) sizes of the same brand, 2) consumer brand-size preference is duration dependent and it first increases and then decreases with the time that consumers stay with a particular brand-size, and 3) consumer price sensitivities are higher for larger size diapers (as consumers learn more about brand qualities when their babies grow older).

Our policy experiment results include the finding that when faced with a price promotion, consumers are more likely to switch from premium (national) brands to economy (store) brands when buying larger sizes than smaller sizes. Our free sampling simulation analysis indicates that while free sampling is overall more beneficial for national brands and providing free samples of smaller sizes is better than providing free samples of larger sizes for all brands, the differential gain between smaller size versus larger size free samples is bigger for store brands.

These combined results suggest a number of managerial implications. First, our results show that consumers who just enter the diaper market are less price sensitive, while they become more price-sensitive when their babies grow (and they gain more experience). The results from our study could help managers do a better job at developing promotion strategies to different consumer segments. More specifically, national brands could focus on providing free samples to consumers who are new to the market, and focus on price promotions to consumers with more category experience. We show also that consumer preference for a brand is duration-dependent and it could decline when their needs change (e.g. the baby is growing out of a size). Managers could develop promotion strategies by tracking a consumer's purchase history and giving free samples (or coupons) around the time that her needs change. Second, given increased price sensitivities over time, store (or private) label managers could determine the optimal time of pursuing a consumer aggressively. National brands, on the other hand, may try to reverse this trend by adding new features to larger size diapers.

Going beyond the diapers category, the general implications of this research include the fact that firms need to be aware of the timing of consumers increased motivation for sampling (e.g., in brand relaunches, the timing will be the same for all households whereas in other cases, specific demographics changes will lead household-level specific timing implications), as well as systematically evolving consumer sensitivities. Varying-parameter models (e.g., Mela, Gupta, Lehmann 1997) have captured stochastically evolving preferences or preferences that evolve as a function of marketing mix and this research shows a systematic evolution of such sensitivities over time in markets where there is quality uncertainty.

References

Ackerberg, D., 2003, Advertising, learning, and consumer choice in experience good markets: A structural empirical examination. *International Economic Review* 44**,** 1007-1040.

Bucklin, Randolph E., and Sunil Gupta., 1992, "Brand Choice, Purchase Incidence, and Segmentation: An Integrated Modeling Approach." *Journal of Marketing Research*, 29 (May 1992): 201-215.

Che, Hai, K. Sudhir and P. B. Seetharaman (2007), "Bounded Rationality in Pricing under State Dependent Demand: Do Firms Look Ahead? How Far Ahead?" *Journal of Marketing Research* 44 (3), 434–49.

Ching, A., T. Erdem and M. Keane, 2012, Learning Models: An Assessment of Progress, Challenges and New developments. *Working Paper*, University of Toronto.

Ching, A., T. Erdem and M. Keane, 2013, A Simple Approach to Estimate the Roles of Learning and Inventories in Consumer Choice. Working Paper, University of Toronto.

Chung, Doug J., Thomas Steenburgh, and K. Sudhir. "Do Bonuses Enhance Sales Productivity? A Dynamic Structural Analysis of Bonus-Based Compensation Plans." *Harvard Business School Working Paper*, No. 13–066, January 2013.

Crawford, G. S. and M. Shum, 2005, Uncertainty and learning in pharmaceutical demand. *Econometrica* 73 (4), 1137-1173.

DeGroot, M. H., 1970, Optimal statistical decisions, McGraw-Hill.

Dickstein, Michael J., 2011, Efficient provision of experience goods: evidence from antidepressant choice. *Working paper*, Harvard Business School.

Dubé, J.P., G. Hitsch, and P. Rossi, 2009, Do switching costs make markets less competitive?, *Journal of Marketing Research* 46(4), 435-445.

Erdem, T., 1998, An Empirical Analysis of Umbrella Branding. *Journal of Marketing Research*, 35 (3), 339-351

Erdem, T. and M. P. Keane, 1996, Decision-making under uncertainty: capturing choice dynamics in turbulent Consumer Goods Markets. *Marketing Science* 15, 1-21.

Gupta, Sachin, Anil Kaul and Dick R. Wittink (1996), "Do Household Scanner Panel Data Provide Representative Inferences from Brand Choices: A Comparison with Store Data," *Journal of Marketing Research*, 33 (November), 383–99.

Hartmann, W.R., 2006, Intertemporal effects of consumption and their implications for demand elasticity estimates. *Quantitative Marketing and Economics* 4(4), 325-349.

Hartmann, W.R., 2010, Demand Estimation with Social Interactions and the Implications for

Targeted Marketing, *Marketing Science*, 29(4), 585-601.

Heckman, J. J., 1981, Heterogeneity and state dependence, in: S. Rosen (Ed.), Studies in labor markets. Chicago University Press, Chicago, pp. 91–139.

Heilman, C. M., D. Bowman, and G. P. Wright, 2000, The evolution of brand preferences and choice behaviors of consumers new to a market. *Journal of Marketing Research* 37 (May), 139-155.

Keane, M. P., 1992, A Note on Identification in the Multinomial Probit Model, *Journal of Business & Economic Statistics*, 10 (2), 193-200.

Keane, M. P. and K. I. Wolpin, 1994, Solution and estimation of dynamic programming models by simulation. *Review of Economics and Statistics* 76(4), 648-672.

Keane, M. P., 1997, Modeling heterogeneity and state dependence in consumer choice behavior. *Journal of Business and Economic Statistics* 15(3), 310-327.

McAlister, L., 1982, A dynamic attribute satiation model of variety seeking behavior. *Journal of Consumer Research* 9, 141-150.

McFadden, D., 1974, Conditional logit analysis of qualitative choice behavior, P. Zarembka (Ed.), *Frontiers of Econometrics*. New York: Academic Press, 105-42.

Mela, C. F., Gupta, S. & Lehmann, D. R. (1997). The long-term impact of promotion and advertising on consumer brand choice. *Journal of Marketing Research*, Vol. 34, 248–261.

Osborne, M., 2011, Consumer learning, switching costs, and heterogeneity: A structural examination. *Quantitative Marketing and Economics*, Volume 9, Number 1, March 2011, pp. 25-70(46)

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

Rust, J., 1987, Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher. *Econometrica*, 55(5), 999-1033.

Shin, S.. S. Misra, and D. Horsky, 2010, "Disentangling preferences and learning in brand choice models. *Marketing Science,* forthcoming.

Sun, B., 2005, Promotion effect on endogenous consumption. *Marketing Science* 24(3), 430-443.

Yao, Song, Carl F. Mela, Jeongwen Chiang, and Yuxin Chen (2012), "Determining Consumers' Discount Rates With Field Studies," Journal of Marketing Research, 49, 6 (December), 822-841.

Table 1: Sales, Revenue, and Market Share Summary Statistics

Brands	Units Sold	Revenue	Purchase Shares
Store	937	9,559.57	0.052
Huggies	4812	87,201.72	0.47
Pampers	4368	88,716.71	0.478

Sizes	Units Sold	Revenue	Purchase Shares
1	1511	22,676.69	0.122
2	1500	28,143.83	0.151
3	2787	51,564.35	0.278
4	2716	53,709.00	0.289
5	1603	29,693.83	0.160

Table 2: Marketing Mix Summary Statistics

Brand	Size	Mean Price	Mean Feature
Store	1	0.151	0.004
Huggies	1	0.223	0.008
Pampers	1	0.232	0.006
Store	2	0.178	0.003
Huggies	2	0.243	0.01
Pampers	2	0.244	0.007
Store	3	0.199	0.005
Huggies	3	0.281	0.009
Pampers	3	0.287	0.008
Store	4	0.22	0.006
Huggies	4	0.347	0.009
Pampers	4	0.314	0.008
Store	5	0.272	0.008
Huggies	5	0.378	0.014
Pampers	5	0.417	0.012

Table 3: Brand-Size Switching Matrix

	1	1	1	2	2	2	3	3	3	4	4	4	5	5	5
Brand	ST	HU	PA	ST	HU	PA	ST	HU	PA	ST	HU	PA	ST	HU	PA
ST	21	5	3	15	2	0	0	0	0	0	0	0	0	0	0
HU	2	63	4	3	33	2	0	0	0	0	0	0	0	0	0
PA	3	3	82	2	3	43	0	0	0	0	0	0	0	0	0
ST	1	0	1	15	1	0	5	1	1	0	0	0	0	0	0
HU	1	9	0	2	92	4	2	36	3	0	0	0	0	0	0
PA	1	0	7	0	2	47	5	3	20	0	0	0	0	0	0
ST	0	0	0	4	1	0	25	2	0	3	1	0	0	0	0
HU	0	0	0	3	25	2	4	132	15	6	54	3	0	0	0
PA	0	0	0	0	0	10	2	17	93	8	4	29	0	0	0
ST	0	0	0	0	0	0	2	5	0	4	4	2	4	1	0
HU	0	0	0	0	0	0	4	37	1	6	96	4	9	30	2
PA	0	0	0	0	0	0	5	2	21	7	7	71	7	4	22
ST	0	0	0	0	0	0	0	0	0	4	0	0	10	5	2
HU	0	0	0	0	0	0	0	0	0	3	19	4	5	83	6
PA	0	0	0	0	0	0	0	0	0	2	2	15	7	7	52

ST=Store, HU=Huggies, PA=Pampers

Table 4: Model Selection

		Estimation Sample		
		1 segment	2 segments	3 segments
Model 1: Learning Model with Myopic Consumers I	LL	-5432.3	-5381.25	-5308.37
	AIC	10906.60	10814.50	10682.74
	BIC	11056.04	10999.52	10917.58
Model 2: Learning Model with Myopic Consumers II (**Model 1** + Learning across Adjacent Sizes)	LL	-5391.43	-5313.94	-5208.3
	AIC	10832.86	10687.88	10490.60
	BIC	11010.77	10901.37	10753.90
Model 3: Learning Model with Myopic Consumers III (**Model 2** + Duration-dependent Parameters in the Utility)	LL	-5300.81	-5339.15	-5220.95
	AIC	10663.62	10752.30	10529.90
	BIC	10884.22	11015.60	10843.01
Model 4: Learning Model with Forward-looking Consumers I (**Model 3** + Forward-looking Consumers)	LL	-5282.08	-5154.71	-5024.93
	AIC	10626.16	10383.42	10137.86
	BIC	10846.76	10646.72	10450.97
Model 5: Learning Model with Myopic Consumers IV (**Model 3** + Size-specific Price Parameters)	LL	-5278.11	-5158.29	-5037.74
	AIC	10622.22	10398.58	10175.48
	BIC	10857.06	10690.35	10531.29
Model 6: Learning Model with Forward-looking Consumers II (**Model 5** + Forward-looking Consumers) as well as (**Model 4**+Size-Specific Price Parameters)	LL	-5108.33	-5033.52	**-4985.75**
	AIC	10282.66	10147.04	**10067.50**
	BIC	10517.50	10431.69	**10409.08**

* Calibration sample: Number of observations = 9798. Number of households = 119. Number of periods = 78.

** Note: AIC=-2*Log-likelihood+2*# of parameters; BIC=-2*Log-likelihood+# of parameters*ln (# of observations).

Table 5(a): Estimates of the Homogeneous Part of Model 5 with 3 Segments

Parameter		Estimates	Std error
Size 1 Quality	Store	**-0.891**	0.16
	Huggies	**0.051**	0.02
	Pampers	**0.045**	0.02
Size 2 Quality	Store	**0.011**	0.01
	Huggies	**0.068**	0.02
	Pampers	**0.084**	0.03
Size 3 Quality	Store	**0.107**	0.05
	Huggies	**0.135**	0.03
	Pampers	**0.165**	0.03
Size 4 Quality	Store	**0.025**	0.01
	Huggies	**0.041**	0.02
	Pampers	**0.025**	0.01
Size 5 Quality	Huggies	**0.077**	0.03
	Pampers	**0.098**	0.04
Duration_size1 (s_1)		0.874	0.831
Duration_size2 (s_2)		**0.742**	0.02
Duration_size3 (s_3)		**0.633**	0.03
Duration_size4 (s_4)		**0.558**	0.04
Duration_size5 (s_5)		**0.302**	0.25
Size Correlation Between 1 and 2 (ρ_{12})		**0.457**	0.05
Size Correlation Between 2 and 3 (ρ_{23})		**0.445**	0.04
Size Correlation Between 3 and 4 (ρ_{34})		**0.459**	0.04
Size Correlation Between 4 and 5 (ρ_{45})		**0.512**	0.05
Standard Deviation of Initial Perception Variability (σ_υ):			
Intercept (λ_0)		**9.865**	0.85
Number of purchases in initialization period (λ_1)		**-1.862**	0.78
Standard Deviation of Experience Variability (σ_ξ)		**1.435**	0.09
Standard Deviation of Experience Variability for across size learning (σ_η)		**14.016**	1.33

Note: Parameters highlighted in bold are 95% statistically significant; parameters highlighted in bold and italic are 90% statistically significant.

Table 5(b): Estimates of the
Heterogeneous Part of Model 5 with 3 Segments

	Segment 1		Segment 2		Segment 3	
Parameter	Estimates	Std error	Estimates	Std error	Estimates	Std error
Price ($\beta_{h\_small\_size}$)	**-0.410**	0.20	**-3.951**	1.53	2.328	1.80
Price ($\beta_{h\_large\_size}$)	**-1.428**	2.10	**-8.587**	1.46	0.117	1.98
Duration (μ_h)	**0.010**	0.00	**0.004**	0.00	**0.269**	0.07
Inventory (γ_h)	**0.034**	0.01	**0.028**	0.01	**0.002**	0.00
Utility Weight (w_h)	**6.931**	1.28	**7.017**	0.98	1	
Risk Coefficient (r_h)	**-0.847**	0.42	**-6.641**	1.16	**-11.257**	1.16
Segment Size Weight	**-1.532**	0.07	**-0.277**	0.02		
Size of Segment 1	22%					
Size of Segment 2	76%					
Size of Segment 3	3%					

Note: Parameters highlighted in bold are 95% statistically significant; parameters highlighted in bold and italic are 90% statistically significant.

Table 6(a)
Effects of Learning Spillover Effects on Choice Probabilities

		Store	Huggies	Pampers
κ_{12}=0	Store	-0.03	0.01	0.05
	Huggies	0.01	-0.08	0.05
	Pampers	0.01	0.04	-0.09
κ_{23}=0	Store	-0.05	0.02	0.02
	Huggies	0.01	-0.05	0.06
	Pampers	0.01	0.04	-0.08
κ_{34}=0	Store	-0.03	0.01	0.02
	Huggies	0.01	-0.07	0.04
	Pampers	0.02	0.05	-0.08
κ_{45}=0	Store	-0.01	0.01	0.01
	Huggies	0.01	-0.05	0.02
	Pampers	0.01	0.03	-0.04

Note: κ's are calculated from Equation (5). Entries in the cell are the choice probabilities calculated from the proposed model by allowing one of the spill-over coefficients (κ's) to be zero.

Table 6(b)
Effects of Correlated Initial Perceptions on Choice Probability

	Store		Huggies		Pampers	
Holding all other ρ's to be 0, when:	size *k*	size *k+1*	size *k*	size *k+1*	size *k*	size *k+1*
$\rho_{12} \neq 0$	0.0002	0.0014	0.0040	0.0005	0.0052	0.0014
$\rho_{23} \neq 0$	0.0033	0.0006	0.0053	0.0011	0.0034	0.0002
$\rho_{34} \neq 0$	0.0003	-0.0001	0.0053	0.0053	0.0059	-0.0001
$\rho_{45} \neq 0$	0.0003	0.0003	0.0046	0.0018	0.0005	0.0041

k=1,2,…,4. Entries in the cell are the choice probabilities calculated from the full models by allowing one of the correlation coefficients ($\rho_{k,k+1}$) to be non-zero, while holding other ρ's to be zeros.

Table 7
Effects of Price Cuts on Choice Probability for Different Brands and Sizes

Cut prices of the small size of the three brands by 10%					
		Change in Choice Probability (in Percentage)			
		Store	Huggies	Pampers	Total
Temporary 10% price cut in	Store	3.37	-0.50	-0.41	0.43
	Huggies	-1.41	3.08	-1.50	1.69
	Pampers	-1.46	-1.42	2.89	1.77
Cut prices of the large size of the three brands by 10%					
		Change in Choice Probability (in Percentage)			
		Store	Huggies	Pampers	Total
Temporary 10% price cut in	Store	4.28	-1.24	-1.27	1.42
	Huggies	-1.76	3.55	-1.52	1.81
	Pampers	-1.86	-1.49	3.30	1.92

Table 8(a)
Observed Purchase Frequency

Brand	Store Brand					Huggies					Pampers					Sum
Period\Size	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	
1	1	2	0	0	0	0	1	2	0	3	0	1	5	2	5	22
2	0	3	2	0	1	1	5	2	9	1	7	4	0	2	2	39
3	0	3	7	0	0	5	3	6	3	4	4	0	1	1	2	39
4	0	0	0	2	0	0	3	4	1	1	1	0	2	3	3	20
5	0	0	5	0	1	2	4	4	3	1	3	1	0	1	3	28
6	1	1	0	0	0	0	0	0	2	0	0	4	1	4	6	19
7	2	4	0	2	4	2	2	2	1	3	1	0	1	2	1	27
8	0	1	0	0	1	0	2	2	1	2	0	1	2	1	3	16
9	0	0	1	0	1	5	2	4	3	0	0	2	3	2	2	25
10	0	1	4	0	0	2	1	4	4	3	0	1	4	5	1	30
11	0	0	1	0	1	0	1	2	3	3	3	1	1	2	3	21
12	1	0	0	0	0	0	2	7	2	1	1	1	2	5	0	22
13	0	0	0	0	0	2	5	3	2	2	1	3	1	5	0	24
14	0	0	0	0	1	1	0	12	1	1	1	3	6	3	3	32
15	0	0	0	0	0	1	6	2	9	5	1	0	5	9	4	42
16	0	0	9	0	0	2	5	6	3	2	0	4	4	1	2	38
17	0	2	3	0	1	3	2	6	9	2	1	3	6	1	0	39
18	0	0	2	2	2	0	6	15	1	3	5	5	1	0	4	46
19	4	0	1	0	0	0	2	14	4	2	1	2	9	5	0	44
20	0	0	0	1	0	1	1	8	5	4	1	0	8	1	2	32
21	0	0	2	2	0	2	1	0	6	2	10	5	5	3	2	40
22	2	3	3	1	5	0	1	10	4	1	1	1	1	10	1	44
23	0	0	4	0	0	1	1	6	3	2	4	3	3	2	2	31
24	0	0	0	0	0	0	3	6	5	3	3	0	5	8	11	44
25	2	0	2	2	5	1	3	2	3	4	1	2	1	1	2	31
Sum	13	20	46	12	23	31	62	129	87	55	50	47	77	79	64	795

Table 8(b)
Purchase Frequency from Baseline Simulation

Brand	Store Brand					Huggies					Pampers					Sum
Period\Size	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	
1	1	2	1	0	0	0	2	2	0	3	1	2	5	2	5	26
2	0	3	1	0	1	1	5	1	7	1	8	4	1	3	2	38
3	0	3	8	0	0	4	4	4	1	4	3	1	1	0	2	35
4	0	0	1	1	0	1	2	4	1	1	1	1	1	3	3	20
5	1	0	5	0	1	3	4	3	3	1	2	2	1	1	3	30
6	0	2	1	0	0	1	2	1	2	0	0	5	0	4	5	23
7	1	2	0	1	3	1	2	2	1	3	2	0	3	3	1	25
8	0	3	0	0	1	2	2	2	1	2	1	1	2	0	3	20
9	0	0	1	0	1	3	3	3	3	0	0	2	3	2	2	23
10	0	1	4	0	0	2	0	5	4	3	0	1	4	4	1	29
11	0	0	1	0	1	0	3	1	3	3	3	2	0	2	4	23
12	0	0	0	0	0	0	2	8	0	1	1	1	1	4	0	18
13	0	0	1	0	0	2	1	2	2	2	1	1	0	6	0	18
14	0	0	0	0	1	1	0	14	1	1	1	4	6	5	3	37
15	0	0	0	0	0	1	4	1	6	5	1	1	5	7	4	35
16	0	0	8	0	0	2	5	7	5	2	0	4	9	1	3	46
17	0	1	4	0	1	3	2	8	8	2	1	4	4	1	0	39
18	2	0	2	3	1	1	6	15	4	3	5	5	1	1	3	52
19	2	0	1	0	0	0	3	12	4	2	1	2	7	4	2	40
20	0	0	1	1	0	0	2	7	2	4	2	0	5	1	1	26
21	0	0	2	1	0	1	1	1	8	2	8	4	6	4	3	41
22	2	1	3	1	4	1	1	11	4	1	3	2	3	9	1	47
23	0	1	2	0	1	1	1	7	3	2	4	3	3	4	2	34
24	0	1	0	0	0	0	2	4	5	3	3	1	5	7	12	43
25	2	0	2	1	3	1	1	4	3	4	1	1	0	2	2	27
Sum	11	20	49	9	19	32	60	129	81	55	53	54	76	80	67	795

			Store					Huggies					Pampers							
Model 6: Learning Model with Forward –looking Consumers II (as reported in Table 5; this model has across-size learning, duration dependence in utility, and size-specific price parameters) LL=-4985.75		**Brand**	**Store**					**Huggies**					**Pampers**					**Change in Own and Adjacent Sizes Between Free Sample and Baseline Simulation**		
		Size	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**			
	Observed		13	20	46	12	23	31	62	129	87	55	50	47	77	79	64			
	Baseline Simulation		11	20	49	9	19	32	60	129	81	55	53	54	76	80	67	Store	Huggies	Pampers
	Free Sample in Size 1-3	Store	18	32	63	12	19	30	52	123	76	55	52	43	73	80	67	36	-21	-15
		Huggies	6	15	43	7	19	42	74	140	80	55	49	46	72	77	66	-18	34	-19
		Pampers	7	12	43	7	19	28	51	125	73	55	67	68	87	85	66	-20	-25	44
	Free Sample in Size 4-5	Store	11	19	51	16	26	32	61	127	77	53	54	54	73	76	65	16	-8	-9
		Huggies	11	19	44	4	13	32	61	137	97	65	53	55	72	69	63	-16	34	-19
		Pampers	11	20	42	4	14	32	60	121	73	50	53	55	85	95	80	-17	-21	37
Model 4: Learning Model with Forward –looking Consumers I (as reported in Table 5; this model has across-size learning, duration dependence in utility, but no size-specific price parameters) LL=-5024.93		**Brand**	**Store**					**Huggies**					**Pampers**					**Change in Own and Adjacent Sizes Between Free Sample and Baseline Simulation**		
		Size	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**			
	Observed		13	20	46	12	23	31	62	129	87	55	50	47	77	79	64			
	Baseline Simulation		11	20	48	10	22	32	61	127	84	54	53	52	76	81	64	Store	Huggies	Pampers
	Free Sample in Size 1-3	Store	24	32	59	12	19	27	52	121	80	55	50	43	73	80	67	38	-24	-16
		Huggies	6	15	43	7	19	42	72	140	87	55	49	46	72	77	66	-18	37	-18
		Pampers	5	10	41	8	19	28	51	123	80	55	67	68	89	84	64	-25	-22	46
	Free Sample in Size 4-5	Store	11	19	51	14	25	32	61	127	78	53	54	53	73	76	64	10	-7	-8
		Huggies	11	19	42	5	11	32	61	131	97	65	53	52	75	71	63	-22	28	-12
		Pampers	11	20	41	4	13	32	60	122	74	52	53	52	84	90	80	-22	-17	33
Model 6a: Model 6 without Duration Dependence in Utility LL=-5003.80		**Brand**	**Store**					**Huggies**					**Pampers**					**Change in Own and Adjacent Sizes Between Free Sample and Baseline Simulation**		
		Size	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**			
	Observed		13	20	46	12	23	31	62	129	87	55	50	47	77	79	64			
	Baseline Simulation		11	21	47	12	20	33	59	129	82	54	51	49	79	80	68	Store	Huggies	Pampers
	Free Sample in Size 1-3	Store	22	32	63	12	19	28	51	121	82	55	47	43	72	80	67	38	-21	-17
		Huggies	6	15	40	12	20	42	74	140	82	55	45	44	72	80	68	-18	35	-18
		Pampers	5	12	38	12	20	28	51	122	82	55	67	68	87	80	67	-24	-20	43
	Free Sample in Size 4-5	Store	11	20	47	20	28	33	59	128	77	53	51	49	78	76	62	16	-7	-11
		Huggies	11	21	47	4	13	32	59	129	97	68	51	55	79	69	63	-15	29	-16
		Pampers	11	20	47	4	11	32	60	129	73	50	51	55	79	95	82	-17	-13	29
Model 6b: Model 6 without Duration Dependence in Utility or Spillover Effect LL=-5012.37		**Brand**	**Store**					**Huggies**					**Pampers**					**Change in Own and Adjacent Sizes Between Free Sample and Baseline Simulation**		
		Size	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**			
	Observed		13	20	46	12	23	31	62	129	87	55	50	47	77	79	64			
	Baseline Simulation		12	20	49	10	20	32	60	130	81	54	51	50	79	78	69	Store	Huggies	Pampers
	Free Sample in Size 1-3	Store	18	33	63	11	19	30	52	123	75	54	52	43	73	78	68	34	-23	-12
		Huggies	6	15	43	10	19	42	75	140	82	55	49	46	72	78	69	-17	36	-13
		Pampers	7	12	41	10	19	27	51	123	77	55	67	68	88	78	69	-21	-25	43
	Free Sample in Size 4-5	Store	12	20	51	17	31	32	61	129	73	52	51	52	78	73	64	20	-11	-11
		Huggies	12	19	48	4	13	32	61	131	98	68	51	50	77	69	63	-14	32	-17
		Pampers	12	20	48	4	12	32	60	128	71	46	51	50	81	97	82	-15	-20	34

Table 8(c) Purchase Frequency from Free Sample Experiments in Different Models