A Dynamic Model of Brand Choice when Price and Advertising Signal Product Quality

Tülin Erdem
Haas School of Business
University of California, Berkeley
E-Mail: erdem@haas.berkeley.edu

Michael P. Keane
Department of Economics
Yale University
E-Mail: michael.keane@yale.edu

Baohong Sun
Carnegie Mellon University, Pittsburgh, PA
E-Mail: bsun@andrew.cmu.edu

February 2005
(Revised February 2006)

Acknowledgements: This research was supported by NSF grant SBR-9812067, "Consumer learning about Quality and its Role in Consumer Choice."

A Dynamic Model of Brand Choice when Price and Advertising Signal Product Quality

Abstract

In this paper, we develop a structural model of household behavior in an environment where there is uncertainty about brand attributes, and both prices and advertising signal brand quality. Four quality signaling mechanisms are at work in the model: 1) price signals quality, 2) advertising frequency signals quality, 3) advertising content provides direct (but noisy) information about quality, and 4) use experience provides direct (but noisy) information about quality. We estimate our proposed model using scanner panel data on ketchup. If price is important as a signal of brand quality, then frequent price promotion may have the unintended consequence of reducing brand equity. We use our estimated model to measure the importance of such effects. Our results imply that price is an important quality signaling mechanism, and that frequent price cuts can have significant adverse effects on brand equity. The role of advertising frequency in signaling quality is also significant, but it is less quantitatively important than price.

Key words: Consumer Choice under Uncertainty, Bayesian Learning, Signaling, Advertising and Price as Signals of Quality, Brand Equity, Pricing Policy, Dynamic Choice

1. Introduction

Consumer learning about quality of alternative brands of an experience good occurs through several channels. In this paper we estimate a dynamic brand choice model in which consumers learn through four channels: use experience signals, advertising content and advertising intensity signals,[1] and price signals. The relative importance of these mechanisms has important implications for how consumer demand responds to changes in price and advertising intensity. Thus, our work is of interest for both marketing and industrial organization.[2]

Prior work has modeled a subset of the quality signaling mechanisms that we consider here. For instance, Erdem and Keane (1996) and Anand and Shachar (2000) estimated models where advertising content and use experience provide noisy signals about brand attributes. In Ackerberg (2003), advertising intensity and use experience both signal product quality. But, to our knowledge, prior empirical work has not incorporated price as a signal of quality in brand choice models. Nor has it allowed simultaneously for the possibilities that advertising may signal quality through both its content and its quantity.

In the theoretical literature, Milgrom and Roberts (1986) developed a model in which price and advertising expenditure signal quality of an experience good.[3] In their model, high quality producers are more likely to enjoy repeat sales. Thus, long-run marginal revenue from advertising (that generates initial sales) is greater for high quality producers. Kihlstrom and Riordan (1984) developed a model in which advertising expenditure signals quality by conveying information about a firm's sunk costs. In their model, high quality raises fixed but not marginal cost. Thus, by spending on advertising, a firm signals to consumers that it thinks it can recover its sunk costs, since its higher product quality will enable it to charge a higher price than low quality firms (that have the same marginal production cost).[4]

---

[1] Throughout this paper, we use the terms “advertising intensity,” “advertising quantity,” and “advertising frequency” interchangeably with advertising expenditure. This is legitimate under the assumption that a brand’s expenditure on advertising determines the frequency with which its ads reach consumers.

[2] Indeed, the recognition that dynamics in consumer demand can have important implications for market equilibrium has recently led to a burst of interest on the part of industrial organization economists in the estimation of dynamic demand models (see, e.g., Ching (2002), Crawford and Shum (2003), Ackerberg (2003), Hendel and Nevo (2002)). Marketers have been interested in the estimation of dynamic demand models using scanner data for many years. Keane (1997) reviews much of this literature. Some more recent work includes Erdem, Imai and Keane (2003) and Mehta, Rajiv and Srinivasan (2003).

[3] Price’s role as a signal of quality has also been discussed by Farell (1980), Gerstner (1985) and Spence (1974).

[4] Price and advertising will function as credible signals only if sellers do not find it profitable to "cheat" by conveying false market signals, for example, charging higher prices for lower quality. Two reasons why sellers might refrain from cheating are desire for repeat sales and presence of informed consumers (Tirole 1991).

These papers were motivated by Nelson (1970), who suggested that much advertising contains no solid content. He argued that firms' advertising expenditures could be rationalized if the volume or intensity of advertising (rather than its content) served as a quality signal in experience goods markets. This view has been challenged by Erdem and Keane (1996), Anand and Shachar (2000) and Erdem and Sun (2002) who argue advertising does convey information.[5] And Resnick and Stern (1977) and Abernethey and Franke (1996), who systematically analyzed TV ads, concluded that the large majority do contain some information content. Thus, it is an empirical question whether advertising signals quality primarily through content or volume.

Similarly, consumers may also use their knowledge of the price-quality relationship that exists in a market to infer quality from price. Price research has shown that the relationship is category specific. For instance, Lichtenstein and Burton (1989) find that both objective quality-price and perceived quality-price relationships are stronger for nondurables. Caves and Greene (1996) found that there is a strong positive relationship between price and (objective) quality for frequently purchased product categories that are convenience goods. Rao and Monroe (1989) argue that a strong positive relationship exists for lower priced, frequently purchased product categories, but that the relationship is not well documented for other categories.

In this paper, we extend the Bayesian learning model of Erdem and Keane (1996) to incorporate both price and advertising frequency as signals of product quality (in addition to use experience and advertising content). Our structural modeling approach will enable us to evaluate the effects of advertising and price promotions both in the short-run and long-run.

A key issue in marketing is whether frequent price promotions or "deals" could have adverse consequences for brand equity (Aaker 1991), i.e., do frequent promotions reduce the perceived quality of a brand, reducing consumer willingness to pay in the long-run? Using a reduced-form model, Jedidi, Mela and Gupta (1999) concluded that advertising increases "brand equity" while promotions reduce it. As our model allows consumers to use price and advertising to signal quality, we will be able to investigate these questions explicitly.

Of course, price fluctuations due to promotions are a salient feature of most frequently purchased consumer goods markets. Consumers in our model use the history of prices to infer the mean price for a brand. It is this mean price that signals brand quality. Thus, if a brand cuts

[5] Anand and Shachar find that increased exposure to ads reduces some consumer's demand for certain brands. The implication is that consumers learn from the ad content that the brand is not a good match with their tastes. If advertising only signals quality through its quantity, then increased exposure would never reduce demand.

its price in week *t*, consumers solve a signal extraction problem to determine the extent to which this represents a transitory fluctuation around the mean vs. a more permanent decline in the brand's mean price. To the extent that consumers revise downward their estimate of the brand's mean price, they will also revise downward their estimate of its quality, reducing brand equity.

Recently, Erdem, Imai and Keane (2003) and Hendel and Nevo (2003) have developed dynamic demand models for frequently purchased storable consumer goods. In these inventory models, consumers attempt to time purchases to occur in periods when prices are relatively low. Thus, both inventory and learning models contain a mechanism whereby, if a brand shifts to a strategy of more frequent "deals," consumer demand for the brand at any given price will fall.

Consistent with his prediction, a substantial reduced form literature in marketing, originating with Winer (1986), has shown that the fit of demand models is substantially improved if they include not just a brand's current price, but also some measure of its "reference price," typically operationalized as an average of lagged prices. The learning model in which price signals quality, and the inventory model, provide alternative rationalizations for reference prices.[6] Thus, an important avenue for future research is to develop methods to distinguish the demand effects of more frequent promotions operating through changes in expected future prices vs. perceived quality. Given current computational technology, it is not now feasible to incorporate both consumer learning about quality and inventory behavior into a single model.

We estimate our model on scanner data from the ketchup category. This may seem unglamorous, but this category is well suited to the investigation. One dominant brand (Heinz) is generally perceived as being high quality. It is also higher priced and has substantially higher advertising expenditure than its name brand competitors, Hunts and Del Monte. In fact, the lowest priced name brand (Del Monte) does not engage in any TV advertising. Thus, there is scope for consumers to use price and ad expenditures as signals of quality in this market.

Our model sheds light on the importance of different information sources in influencing perceived quality. For instance, it implies that price does play a significant role. We predict that a 10% permanent price cut for Heinz would increase its sales by 26%. However, if the price cut could be implemented without reducing perceived quality (and, hence, brand equity), we predict that the increase in sales would be much greater (32%).

---

[6] Another type of evidence used to support inventory models is that duration to next purchase tends to be longer after a deal purchase. This could also be rationalized by a learning model with an outside good, if the deal causes quality perceptions to be revised downward.

II. The Model

II.1. Overview

We model household behavior in an environment where households have uncertainty about quality levels of brands, and may be risk averse with respect to quality variation. Households may use prices, use experience, advertising frequency and advertising content as signals of brand quality. They use the frequency with which they see TV ads for a brand as a signal of that brand's level of advertising expenditures.[7] Households update their expectations of brand quality in a Bayesian manner as they see additional signals.

We do not attempt to model producer behavior. Rather we specify functional relationships between price, advertising frequency and quality that we assume hold in equilibrium. We estimate the parameters of these functional relationships jointly with the parameters of our structural model of household behavior. Households are assumed to know these equilibrium relationships, and to use them to help infer brand quality.[8]

We estimate a pure brand choice model, ignoring the issues of quantity choice and inventories that are the focus of Erdem, Imai and Keane (2003) and Hendel and Nevo (2003). Those papers ignore consumer learning. Thus, each approach leaves out a potentially important aspect of consumer behavior. A unification of these two approaches is left for future research.

II.2. Utility Function

We assume consumers have utility functions of the form:

$$U_{ijt} = \alpha_i P_{ijt} + w_i Q_{Eijt} + w_i r_i Q_{Eijt}^2 + e_{ijt} \qquad (1)$$

where P_{ijt} is the price of brand $j=1,\ldots,J$ faced by household i at time t, and Q_{Eijt} is household i's *experienced* quality of brand j at time t. The parameter α_i, the price coefficient, is the negative of household i's marginal utility of consumption for the outside good. It is assumed constant over the small range of outside good consumption levels generated by the household's brand choice decisions. The parameter w_i is utility weight that household i places on quality. The parameter r_i

[7] In most mass marketed and frequently purchased product categories, brands are nationally advertised on TV. Air time is priced based on viewership (i.e., ratings). Thus, it is reasonable to assume that the frequency with which a person sees ads for a brand is proportional to that brand's advertising expenditure. Implicitly, we assume that if a brand buys a time slot with twice the audience rating, the cost of the slot is doubled and the probability a randomly drawn consumer is exposed to the ad also doubles. Thus, cost per ad view is assumed constant.

[8] Such a quasi-structural approach to estimation has been adopted previously in various differing contexts by Ching (2002), Erdem, Imai and Keane (2003) and Hendel and Nevo (2003), who estimate reduced form pricing policy functions jointly with various structural models of consumer choice behavior.

captures household i's degree of risk aversion regarding variation in quality. Finally, e_{ijt} is a preference shock known to the household but unobserved by the econometrician.

Variability of *experienced* quality Q_{Eijt} around a brand's true quality Q_j occurs for several reasons. One is variability of product quality across units. But, in categories covered by scanner data, a more plausible explanation is that a user's experience of a brand is context dependent. Thus, we assume that each use experience provides a noisy but unbiased signal of quality, according to $Q_{Eijt}=Q_j+\xi_{ijt}$ where $\xi_{ijt}\sim N(0,\sigma_\xi^2)$. We refer to σ_ξ^2 as the "experience variability."

Household i has an information set I_{it} containing all brand quality signals it has received up through time t. Given this information, it forms an expectation of Q_{Eijt}. Let $Q_{ijt}\equiv E[Q_j|I_{it}]$ denote household i's expectation of brand j's true quality level at time t. We describe the contents of I_{it} and how expectations are formed below. For now, we just note that all signals are assumed unbiased. Hence, $E[Q_{Eijt}|I_{it}]=E[Q_j|I_{it}]=Q_{ijt}$, and we may write $Q_{Eijt}=Q_{ijt}+(Q_j-Q_{ijt})+\xi_{ijt}$.

Hence, the expected utility to household i from buying and consuming brand j at time t is:

(2) $$E[U_{ijt}|I_{it}]=\alpha_i P_{ijt}+w_i Q_{ijt}+w_i r_i Q_{ijt}^2+w_i r_i E[(Q_j-Q_{ijt})^2\,|\,I_{it}]+w_i r_i\sigma_\xi^2+e_{ijt}$$

In (2), there are two sources of *expected* variability of experienced quality Q_{Eijt} about true quality Q_j. The first is experience variability, captured by σ_ξ^2. The second is $E[(Q_j-Q_{ijt})^2\,|\,I_{it}]$, the variability of true quality around perceived quality; i.e., the household understands it has incomplete information, and that the true quality of a brand will, in general, depart somewhat from its expectation. If a household has little information about a brand, this "risk term" will tend to be large. Thus, *ceteris paribus*, risk averse households will tend to avoid an unfamiliar brand in favor of a familiar brand, even if both brands have the same expected quality.

Note that equation (1) is the same type of utility function used by Erdem and Keane (1996). However, unlike Erdem and Keane, we let the parameters α_i, w_i and r_i be heterogeneous across consumers. We adopt a discrete mass point (latent class) approach to modeling heterogeneity, as in Heckman and Singer (1981).[9] We thus estimate a vector (α_k, w_k, r_k) for each segment of consumers $k=1, \ldots, K$, as well as the population type proportions for each segment, which we denote by π_k for $k=1, \ldots, K$.

[9] We must solve a household's dynamic optimization problem conditional on it being each of the K types in order to form the likelihood function of our model. Thus, it is infeasible to have a continuous heterogeneity distribution.

II.3. The Price Process and the Price-Quality Relation

In using prices to infer quality, consumers assume that the stochastic process for prices is:

(3) $ln\ P_{ijt} = P_j^M + \omega_{ijt},$ $\qquad \omega_{ijt} \sim N(0,\ \sigma_\omega^2)$

where P_{ijt} is the price of brand j faced by household i at time t, P_j^M is the mean of the log price of brand j, and ω_{ijt} is a stochastic term that is i.i.d. over time.[10] Consumers believe that, in the market equilibrium, the mean price P_j^M is related to brand quality according to the relation:

(4) $P_j^M = P_0 + \phi Q_j + \eta_j$

where Q_j is a latent quality index for brand j, ϕ is a parameter, and η_j represents the deviation of brand j from the "typical" price quality relationship (i.e., some brands may have prices that tend to be high or low relative to their quality level).

Households perceive that the η_j are distributed in the population of firms according to:

(5) $\eta_j \sim N(0, \sigma_\eta^2)$

Combining equations (3) and (4) we have:

(6) $ln\ P_{ijt} = P_0 + \phi Q_j + \eta_j + \omega_{ijt}$

We will estimate P_0, ϕ, σ_ω, σ_η and a set of η_j. Obviously we cannot estimate both P_0 and a value of η_j for each brand, so we restrict $\eta_J = \sum_{j=1,J-1} \eta_j$ so that the η_j are mean zero across brands.

II.4. Consumer Learning About Quality: The Case of Price as the Only Signal

To illustrate how households learn about quality in our model, it is helpful to consider a hypothetical case where price is the only signal. At $t=0$, prior to any experience in the market, a household has priors about the mean prices and quality levels of brands. The prior for quality is:

(7) $Q_j \sim N(Q_0, \sigma_{Q_0}^2)$ $\qquad$ for $\qquad j = 1,\ldots, J$

and, combining (4), (5) and (7), the prior for mean log price is:

(8) $P_j^M \sim N(P_0 + \phi Q_0,\ \phi^2 \sigma_{Q_0}^2 + \sigma_\eta^2)$ $\qquad$ for $\quad j = 1,\ldots, J.$

In (7) the household's prior is that all brands have a quality level of Q_0, but that the true quality of brand j has variance $\sigma_{Q_0}^2$ around that mean. The prior perceived standard deviation σ_{Q_0} is a

[10] As discussed in Erdem, Imai and Keane (2003) and Hong, MacAfee and Nayyar (2002), prices for frequently purchased consumer goods exhibit complex serial correlation patterns. However, since expected future prices do not affect current consumer demand in our model (as there are no inventories), the consumer behavior predicted by our model should not be too sensitive to whether our assumed price process captures this serial correlation.

parameter to be estimated in our model. The prior mean Q_0 is restricted to equal the mean of the brand specific quality levels Q_j for j=1, …, J, and it is the latter that are estimated.

In (8) the household's prior is that all brands have a mean log price equal to the mean log price in the category, $P_0+\phi Q_0$, but that the true mean log price for brand j has a variance of $\phi^2\sigma^2_{Q0}+\sigma^2_\eta$ around that mean. Note that a brand may have an above average price because it is high quality (the $\phi^2\sigma^2_{Q_0}$ component) or because it is priced high given quality (the σ^2_η component).

Let P^M_{ijt} and Q_{ijt} denote household i's prior means for mean log price and quality of brand j conditional on information at t. At t=0, these are simply $P^M_{ij0}=P_0+\phi Q_0$ and $Q_{ij0}=Q_0$. When a price is observed for brand j at t=1, the household updates its priors about mean log price and quality of brand j using standard Bayesian updating rules (see, e.g., DeGroot (1970)):

(9) $$P^M_{ij1}=P^M_{ij0}+[ln P_{ij1}-P^M_{ij0}]\cdot K^P_{1ij}$$

(10) $$Q_{ij1}=Q_{ij0}+[ln P_{ij1}-P^M_{ij0}]\cdot K^{PQ}_{1ij}$$

where K^P_{1ij} and K^{PQ}_{1ij} are the Kalman gain coefficients, which at $t=1$ are:

(11) $$K^P_{1ij} = (\phi^2\sigma^2_{Q_0}+\sigma^2_\eta)/(\phi^2\sigma^2_{Q_0}+\sigma^2_\eta+\sigma^2_w)$$

(12) $$K^{PQ}_{1ij} = \phi\sigma^2_{Q_0}/(\phi^2\sigma^2_{Q_0}+\sigma^2_\eta+\sigma^2_\omega)$$

Of particular interest is K^{PQ}_{1ij}, which shows how the household revises its perception of the quality of brand j in response to the price surprise. The numerator in (12) is ϕ times the part of perceived price variability that arises because brand j may be of above of below average quality. If there is a positive price surprise (i.e., $lnP_{ij1}>P^M_{ij0}$) the household will revise upward its perception of the quality of brand j provided that $\phi>0$ (i.e., price is related to quality), and provided that $\sigma^2_{Q_0}>0$ (i.e., the household is uncertain about the quality of brand j).[11]

Prior uncertainty about quality σ_{Q_0} is a parameter to be estimated in our model. The size of this parameter determines the extent to which households learn about quality through prices

[11] The denominator in (12) includes all three sources of variation in price. The extent to which households revise quality perceptions in response to a price surprise is inversely related to σ^2_η the variability of mean price not attributable to quality, and σ^2_ω, the variability of prices (about the mean) over time. If these terms are large relative to $\phi^2\sigma^2_{Q_0}$, then most price variability is idiosyncratic and conveys little information about quality.

and other signals. Intuitively, σ_{Q_0} is identified from how brand choice behavior of households with substantial prior experience differs from that of households with little prior experience. If we estimate σ_{Q_0} =0, our model reduces to a static model in which no learning occurs.

As households acquire information, priors become tighter. Let $\sigma^2_{Pijt} = Var(P^M_{ijt} - P^M_j)$ and $\sigma^2_{Qijt} = Var(Q_{ijt} - Q_j)$ denote the household's perceived variability of price and quality for brand j conditional on information received up through time t. At t=0, these perception variances are $\sigma^2_{Pij0} = \phi^2\sigma^2_{Q0} + \sigma^2_\eta$ and $\sigma^2_{Qij0} = \sigma^2_{Q_0}$. Given the price for brand j at t=1, the household updates these prior variances using standard Bayesian updating rules (see, e.g., DeGroot (1970)), to obtain:

(13) $$\sigma^2_{Pij1} = [1/\sigma^2_{Pij0} + 1/\sigma^2_\omega]^{-1}$$

(14) $$\sigma^2_{Qij1} = [1/\sigma^2_{Qij0} + \phi^2/(\sigma^2_\eta + \sigma^2_\omega)]^{-1}$$

According to (13), if σ^2_ω is large then one price signal is not very informative about mean price, so it causes little reduction in perceived variance. Similarly, (14) says that if $\sigma^2_\omega + \sigma^2_\eta$ is large or ϕ is small then a single price realization is not very informative about brand quality.

In period t=2, updating is done in the same way, using the updated means in (9)-(10) as the new prior means, and using the updated variances in (13)-(14) as the new prior variances. Finally, note that for $t \geq 2$ the Kalman gain coefficients for brand j are:

(11') $$K^P_{tij} = \sigma^2_{Pij,t-1} / (\sigma^2_{Pij,t-1} + \sigma^2_\omega)$$

(12') $$K^{PQ}_{tij} = \phi\sigma^2_{Qij,t-1} / (\phi^2\sigma^2_{Qij,t-1} + \sigma^2_\eta + \sigma^2_\omega)$$

II.5. Introducing Advertising Frequency as a Signal of Quality

Now we also incorporate advertising frequency as a signal of brand quality. Let A_{ijt} denote the (normalized) number of TV ads seen by household i for brand j during week t. The normalization is needed to adjust for how often a household watches TV (i.e., how often it sees ads in general). Thus, we normalize A_{ijt} as follows: First, we find the mean number of ketchup ads (for all brands) that a household sees per week during the entire sample period. Second, we scaled this variable so it has a mean of one. We then use this variable to normalize A_{ijt}.

The distribution of A_{ijt} is highly non-normal due to the concentration of mass at zero ads. Unfortunately, we cannot allow for non-normal errors because of the great difficulty of implementing Bayesian updating rules with multiple signals if some are non-normal. Thus, we use a Box-Cox transformation to bring the ad exposure distribution more closely into line with normality. As the Box-Cox likelihood is not well behaved when the dependent variable is zero, we use $1+A_{ijt}$ rather than A_{ijt} itself. Thus, the assumed stochastic process for ad exposures is:

(15) $$[(1+A_{ijt})^{\beta} - 1]/\beta = A_j^M + \theta_{ijt} \quad \text{with} \qquad \theta_{ijt} \sim N(0,\sigma_{\theta}^2)$$

where β is the Box-Cox parameter, A_j^M is the mean of transformed weekly advertising exposures for brand j, and θ_{ijt} is a stochastic term that is i.i.d. over time.[12] The stochastic term captures the idiosyncratic reasons that a household might see more or fewer ads than usual for brand j during week t. These include that the brand's true ad intensity varies by week, along with the luck of the draw (i.e., whether the household happens to be watching TV when the brand's ads appear).

Households believe that A_j^M is related to brand quality according to the relation:

(16) $$A_j^M = A_0 + \delta Q_j + \mu_j$$

Here, Q_j is the latent quality index for brand j, which also appeared in (4), δ is a parameter and μ_j represents the departure of brand j from the "typical" ad frequency-quality relationship (i.e., some brands may advertise relatively heavily given their quality level).

Households perceive that the μ_j are distributed in the population of firms according to:

(17) $$\mu_j \sim N(0,\sigma_{\mu}^2).$$

Combining (15) and (16), we have:

(18) $$[(1+A_{ijt})^{\beta} - 1]/\beta = A_0 + \delta Q_j + \mu_j + \theta_{ijt}$$

We will estimate A_0, β, δ, σ_θ, σ_μ and a set of μ_j. Obviously, we cannot estimate both A_0 and a value of μ_j for each brand, so we restrict $\mu_J = \sum_{j=1,J-1} \mu_j$ so the μ_j are mean zero across brands.

At t=0, prior to seeing any ads, households' prior is that each brand j's (transformed) advertising rate is the same as the mean rate in the category, $A_0+\delta Q_0$, but that the true rate for brand j is distributed around that mean according to:

(19) $$A_j^M \sim N(A_0+\delta Q_0, \delta^2 \sigma_{Q_0}^2 + \sigma_{\mu}^2)$$

Note that a brand may have an above average advertising rate due to either high quality or the deviation (μ_j) from the "typical" ad frequency-quality relationship.

[12] The right side of (18) must be greater than $-1/\beta$ for $(1+A_{ijt})^{\beta} > 0$, which is necessary for the implied value of A_{ijt} to be well defined (given $0<\beta<1$). Thus, taken literally, (18) rules out normal errors. This is an oft-noted problem with Box-Cox transformations. Given our estimates, a value of the right side less than $-1/\beta$ is an extreme outlier.

Let A^M_{ijt} denote household i's prior mean for (transformed) advertising frequency of brand j conditional on information at t. At t=0, this is simply $A^M_{ij0}=A_0+\delta Q_0$. The formulas for how the household updates its perceptions of A^M_{ijt} and Q_{ijt} based on observing a certain (normalized) number of ads A_{ijt} are exactly analogous to equations (9)-(14) and are given by:

(20) $$A^M_{ij1}=A^M_{ij0}+[((1+A_{ij1})^{\beta}-1)/\beta-A^M_{ij0}]\cdot K^A_{ij1}$$

(21) $$Q_{ij1}=Q_{ij0}+[((1+A_{ij1})^{\beta}-1)/\beta-A^M_{ij0}]\cdot K^{AQ}_{ij1}$$

(22) $$K^A_{ij1}=(\delta^2\sigma^2_{Q_0}+\sigma^2_{\mu})/(\delta^2\sigma^2_{Q_0}+\sigma^2_{\mu}+\sigma^2_{\theta})$$

(23) $$K^{AQ}_{ij1}=\delta\sigma^2_{Q_0}/(\delta^2\sigma^2_{Q_0}+\sigma^2_{\mu}+\sigma^2_{\theta})$$

(24) $$\sigma^2_{Aij1}=[1/\sigma^2_{Aij0}+1/\sigma^2_{\theta}]^{-1}$$

(25) $$\sigma^2_{Qij1}=[1/\sigma^2_{Qij0}+\delta^2/(\sigma^2_{\mu}+\sigma^2_{\theta})]^{-1}$$

where the equations for $t\geq 2$ are analogous to equations (11') and (12').

Notice that updating of quality perceptions for brand j based on realized ad exposures given in equation (21) induces an update in the prior for mean log price, provided a price-quality relationship exists in the market. Thus, ad exposures and price realizations interact in the updating of mean price, ad frequency and quality perceptions. The effect of a series of price and ad exposure realizations on a household's perceptions about a brand can be obtained by stringing together updating equations of the form of (9)-(10), (13)-(14), (20)-(21) and (24)-(25).

II.6. Introducing Use Experience and Advertising Content Signals

Use experience and advertising content also provide direct but noisy information about product quality, as in Erdem and Keane (1996). Define d_{ijt} as an indicator equal to 1 if brand j is purchased at time t and 0 otherwise. As we noted in section II.2, use experience provides a direct but noisy information signal Q_{Eijt} according to:

(26) $$Q_{Eijt}=Q_j+\xi_{ijt} \qquad \text{with} \qquad \xi_{ijt}\sim N(0,\sigma^2_{\xi}).$$

Advertising exposure provides a direct but noisy information signal AD_{ijt} according to:

(27) $$AD_{ijt}=Q_j+\tau_{ijt} \qquad \text{with} \qquad \tau_{ijt}\sim N(0,\sigma^2_{\tau}).$$

The updating of expectations with use experience and ad content signals is described in detail in Erdem and Keane (1996), so we will not repeat that here.

II.7. The Household's Dynamic Optimization Problem

The state of a household at time t is characterized by its time t priors for mean log prices, advertising frequencies, and quality levels for all brands, as well as its perception error variances. Let I_{it} denote the state of household i at the point in period t when ads and prices have been observed, but before the purchase decision has been made. We have:

$$I_{it}=[\{Q_{ijt},P^M_{ijt},A^M_{ijt},\sigma^2_{Qijt},\sigma^2_{Pijt},\sigma^2_{Aijt}\},j=1,J].$$

Then the expected time t utility conditional on choosing alternative j in state I_{it} is:

$$\begin{aligned}\mathrm{E}[U_{ijt}\,|I_{it}]&=\alpha_i P_{ijt}+w_i\mathrm{E}[Q_{Eijt}\,|I_{it}]+w_ir_i\mathrm{E}[Q^2_{Eijt}\,|I_{it}]+e_{ijt}\\&=\alpha_i P_{ijt}+w_i\mathrm{E}[Q_{Eijt}\,|I_{it}]+w_ir_i\mathrm{E}[Q_{Eijt}\,|I_{it}]^2+w_ir_i\sigma^2_{Qijt}+w_ir_i\sigma^2_{\xi}+e_{ijt}\\&=\alpha_i P_{ijt}+w_iQ_{ijt}+w_ir_iQ^2_{ijt}+w_ir_i\sigma^2_{Qijt}+w_ir_i\sigma^2_{\xi}+e_{ijt}\end{aligned}$$

We specify that the expected utility from no purchase, $\mathrm{E}[U_{i0t}|I_{it}]$, is given by:

$$\mathrm{E}[U_{i0t}\,|I_{it}]=\Phi_0+\Phi_1\cdot t+e_{i0t}$$

The time trend in this equation captures changes in the value of the outside option over time.

In our model, a household's time t purchase decision affects not only its time t utility, but also its state $I_{i,t+1}$ at the start of period $t+1$. Hence, if a household has little information about a brand, it may be optimal to try it when it is on sale, as it may be better than the household's "preferred" brand (i.e., that with highest expected utility given current information).

The household's optimal decision rule is to choose the option that maximizes the expected present value of utility over the planning horizon. This leads to a dynamic programming (DP) problem. We apply Bellman's principle to solve this problem by finding value functions corresponding to each alternative choice. Letting λ denote the discount factor,[13] the value of choosing alternative j at time t is:

(28) $$V_{ijt}(I_{it})=E[U_{ijt}\,|I_{it}]\;+\;\lambda\,E[\max_l V_{il}(I_{i,t+1})\,|\,I_{it}]\text{ for }\quad j=0,J.$$

We will assume that the stochastic terms $\{e_{ijt}\}_{j=1,J}$ and e_{i0t} that enter the $E[U_{ijt}\,|I_{it}]$ terms are distributed i.i.d. extreme value. We can then obtain closed form expressions for the Emax functions (see Rust (1987)), *conditional* on the state I_{it}.

[13] We fixed the weekly discount factor λ at 0.995, rather than estimating it, because 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.

We solve the DP problem via "backsolving" from a terminal period T. In the process, we need to calculate the Emax functions $E[\max_k V_{ik,T-1}(I_{i,T-1}) \mid I_{i,T-2}]$ at every possible state in I_{it}. A severe complication arises because the set of points in I_{it} is infinite, since there are six continuous state variables. To deal with this problem, we solve for the Emax functions at only a randomly selected finite number of state points, and use an interpolating regression in the state variables to find Emax at other points.[14] We set T=100, which is 49 weeks past the last observation period.[15]

II.8. Constructing the Likelihood Function

Let Θ_k denote the complete set of model parameters for a household of type k, and define:

$$V_{ijt}^{*}(I_{it} \mid \Theta_k) = V_{ijt}(I_{it} \mid \Theta_k) - e_{ijt}$$

so V^* is the deterministic part of the value function in (28). Then we have:

$$Prob\ (\ d_{ijt}=1 \mid I_{it}, \Theta_k) = \exp\{V_{ijt}^{*}(I_{it} \mid \Theta_k)\} / \Sigma_{m=0,J} \exp\{V_{imt}^{*}(I_{it} \mid \Theta_k)\}$$

Let $H_i = \{\{d_{ijt}\}_{j=0}^{J}\}_{t=1}^{T^*}\}$ denote i's choice history, where T^* is the last observation period. Then:

$$Prob(\ H_i \mid \Theta_k) = \prod_{t=1}^{T^*} \Pr ob(d_{ijt} = 1 \mid I_{it}, \Theta_k)^{d_{ijt}}$$

Next, let $\tilde{\xi}_{it} = \{\{d_{ijt}\xi_{ijt}\}_{j=1}^{J}\}_{t=1}^{t}\}$ and $\tilde{\tau}_{it} = \{\{D_{ijt}\tau_{ijt}\}_{j=1}^{J}\}_{t=1}^{t}\}$ denote the sets of use experience and ad content signals, respectively, received by household i up through time t, so that $I_{it} = I_{it}(\tilde{\xi}_{it}, \tilde{\tau}_{it})$. Then, we can write the probability of the observed history for household i as:

$$Prob(\ H_i \mid \Theta_k) = \int_{\tilde{\xi}_{iT^*}} \int_{\tilde{\tau}_{iT^*}} \prod_{t=1}^{T^*} \Pr ob(d_{ijt} = 1 \mid I_{it}(\tilde{\xi}_{it}, \tilde{\tau}_{it}), \Theta_k)^{d_{ijt}}\, dF(\tilde{\xi}_{iT^*}, \tilde{\tau}_{iT^*})$$

We simulate this high dimensional integral using draws for $\tilde{\xi}_{iT^*}$ and $\tilde{\tau}_{iT^*}$ Letting $(\tilde{\xi}_{iT^*}^{m}, \tilde{\tau}_{iT^*}^{m})$ denote the mth draw for household i, where m=1,…, M, we have the unbiased and consistent simulator:

$$\hat{\Pr ob}(\mathrm{H_i} \mid \Theta_k) = \sum_{m=1}^{M} \prod_{t=1}^{T^*} \Pr ob(d_{ijt} = 1 \mid I_{it}(\tilde{\xi}_{it}^{m}, \tilde{\tau}_{it}^{m}), \Theta_k)^{d_{ijt}} \tag{29}$$

We set simulation size M equal to 100, and found that results we not sensitive to increases in M.

Next, define the residuals in the price-quality and advertising-quality relationships as ω_{ijt}

[14] This approximation procedure was developed in Keane and Wolpin (1994). The explanatory variables we use are a constant, and for each of the three brands, the perception error standard deviations for quality, mean price, and mean ad frequency, as well as expected quality, price, and ad frequency.

[15] The interpolating coefficients stabilized before we backsolved to t=51, and were insensitive to increasing T.

$= ln\ P_{ijt}$ - P_0 - $\phi\, Q_j$ - η_j and $\theta_{ijt} = [(1+A_{ijt})^{\beta}$ - $1]/\beta$ - A_0 - $\delta\, Q_j$ - μ_j. Letting $f(\omega)$ and $f(\theta)$ denote the densities of ω and θ, the simulated likelihood for household i is:

(30) $$L_i(\Theta) = \prod_{t=1}^{T^*}\prod_{j=1}^{J} P_{ijt}^{-1} f(\omega_{it})(1+A_{ijt})^{\beta-1} f(\theta_{it}) \sum_k \pi_k \,\hat{\text{Prob}}\,(\text{H}_\text{i} \mid \Theta_k)$$

Note that the Jacobian, $|d\theta/dA|$, generates the term $(1+A_{ijt})^{\beta\text{-}1}$ in the likelihood.

II.9. The Initial Conditions Problem

The first observation period does not coincide with the start of a household's choice process, creating an initial conditions problem. Since our data only contain ad viewing data for the last 51 weeks, we use the first 102 weeks to estimate each household's initial conditions, and the last 51 to estimate the model. Assume household i's prior variance on the quality level of brand j at the start of our estimation period is given by:

(31) $$\ell n \sigma_{Q_{ij0}} = k_0 - k_1 \sum_{\tau=-101}^{0} d_{ij\tau}$$

where k_0 and k_1 are parameters. (31) says initial uncertainty about brand j is less if a household bought j more during the proceeding 102 weeks, reducing its prior variance on j from $\sigma_{Q_0}^2$ to $\sigma_{Q_{ij0}}^2$. This is equivalent to the variance reduction from a single hypothetically quality signal with variance equal to $\sigma_{xij}^2 = [1/\sigma_{Q_{ij0}}^2 - 1/\sigma_{Q_0}^2]^{-1}$.

Thus, we integrate over initial conditions as follows: For each household i, we draw a set of hypothetical signals $\{\{x_{ij}^m\}_{j=1}^J\}_{m=1}^M\}$ from the distribution $x_{ij} \sim N(Q_j, \sigma_{xij}^2)$. Denote by $\tilde{x}_i^m$ the vector of signals for all brands received by i under draw m. When simulating the likelihood contribution for household i, we append this draw to the mth draw for the in-sample use experience and ad content signals, and obtain:

(32) $$\hat{\text{Prob}}(\text{H}_\text{i} \mid \Theta_k) = \sum_{m=1}^{M}\prod_{t=1}^{T^*} \Pr ob(d_{ijt} = 1 \mid I_{it}(\tilde{\xi}_{it}^m, \tilde{\tau}_{it}^m, \tilde{x}_i^m), \Theta_k)^{d_{ijt}}$$

II.10. Identification

Discrete choice models typically require scale and location normalizations on utility for identification. One can scale all the Q_j by a positive constant λ, while scaling σ_ξ, σ_τ and σ_{Q_0} by λ, and w, r, ϕ, and δ by a factor of λ^{-1}, leaving behavior implied by the model unchanged. Thus, we set Q_1 =1 (i.e., Heinz quality = 1), and measure quality of other brands relative to Heinz.

Next, consider the location normalization. First, consider the subset of households with sufficient prior experience of all brands that $\sigma_{Q_0} = 0$. Such households have no uncertainty about brand attributes, so, for them, our model reduces to a static model with no learning. For these households, $Q_{ijt} = Q_j \ \forall \ i,j$, and (ignoring heterogeneity in parameters) equation (2) reduces to:

(2') $$E[U_{ijt} | I_{it}] = \alpha P_{ijt} + wQ_j + wrQ_j^2 + wr\sigma_\xi^2 + e_{ijt}$$

Thus, the "alternative specific intercepts" are $w + wr + wr\sigma_\xi^2$ for j=1, $wQ_j + wrQ_j^2 + wr\sigma_\xi^2$ for j=2, …, J, and Φ_0 for the outside alternative. Obviously this model is not identified, without a location normalization like $\Phi_0 = 0$. Then, conditional on prices, the log odds ratios between all pairs of alternatives are determined by J constants $\{k_j\}_{j=1}^J$ given by $k_j = wQ_j + wrQ_j^2 + wr\sigma_\xi^2$. These J equations contain J+2 unknowns, so identification requires two normalizations. For instance, we could set $r = 0$ and $\sigma_\xi = 0$. Then, α, w, and Q_j for $j \geq 2$ are identified in (2'). It is not surprising that the parameters that measure risk aversion and experience variability become unidentified in the static model in which households have no uncertainty about brand quality.

From this discussion, it is apparent that r, σ_ξ and σ_{Q_0} are only identified from dynamics of the model, i.e., how households' brand choice probabilities evolve over time as they receive signals. Consider the case where households only learn about brands via use experience signals. Suppose first that $r = 0$, so households are risk neutral. Then, holding prices fixed, a household's probability of buying a brand will tend to be increasing (decreasing) in its number of prior purchases of that brand if the brand is of above (below) average quality.

However, a pervasive feature of frequently purchased consumer goods is that brand choice probabilities are increasing in number of past purchases for *all* brands, even those of below average quality. Such a pattern can be generated by $r < 0$ or taste heterogeneity. The feature of the data that distinguishes these two stories is whether brand choice probabilities are nonstationary at the household level: That is, if $r < 0$, a household's willingness to pay for a brand is increasing in its prior experience with that brand, even holding perceived quality fixed.

To summarize, in our data there will be households with sufficient experience with all brands that their learning has ceased and their choice behavior is stationary. The parameters α, w, $\{k_j\}_{j=1}^J$, are identified from choice behavior of these households. This leaves the dynamics of choice behavior amongst households with little prior experience to pin down r, σ_{Q_0} and σ_ξ.

Notably, if $r<0$, $\sigma_{Q_0}>0$ and $\sigma_\xi>0$, so that learning occurs, the location normalization $\Phi_0=0$ needed in a static model is no longer required. Consider a shift in brand 1's intercept, induced by shifting w by the increment $\Delta>0$. Is there a transformation of the remaining model parameters that: (1) does not alter choice behavior for households with complete information (i.e., that shifts the alternative specific intercepts $k_j = wQ_j + wrQ_j^2 + wr\sigma_\xi^2$ for $j\geq 1$ and $k_0=\Phi_0$ by equal increments), and: (2) does not alter learning behavior or change attitudes towards risk? If not, then w is identified without normalizing Φ_0.

Increasing w to $w' = w+\Delta$ scales up households' utility weight on quality. Thus, to hold behavior fixed, the scale of the quality measures Q_j for $j\geq 2$ must be compressed towards the base of $Q_1=1$. In the limit, as $\Delta\to\infty$, we must have $Q_j\to 1$ for $j\geq 2$, while Q_1 is unchanged at 1. Thus, without loss of generality, we can write that each Q_j must be transformed according to $Q'_j = \lambda Q_j + (1-\lambda)$, where $\lambda\in(0,1)$ is a parameter that *may differ* across brands j.[16]

Let us assume for the moment that λ does not differ across brands. If $w' = w+\Delta$ then we can show that, in order to hold behavior towards risk fixed, we must use the transformation $r' = r/[\lambda - 2r(1-\lambda)]$. This leaves the peak of the quadratic utility function, and the coefficients of absolute and relative risk aversion, unchanged in terms of the original quality scale Q. Note that, in order to hold behavior towards risk fixed, λ must be <u>common</u> across alternatives j. We now investigate the form of λ. Note that the transformed brand specific intercept for brand 1 is:

$$k'_1 = (w+\Delta) + (w+\Delta)r' + (w+\Delta)r'(\sigma'_\xi)^2$$

while the transformed intercept for any brand $j\geq 2$ is

$$k'_j = (w+\Delta)Q'_j + (w+\Delta)r'(Q')_j^2 + (w+\Delta)r'(\sigma'_\xi)^2.$$

In order for the transformation to leave choice probabilities unchanged for those households with complete information about quality, we require that $k'_j - k_j = k'_1 - k_1$ for $j\geq 2$. This means that λ must solve the following quadratic equation:

$$\begin{aligned}(33)\quad & (w+\Delta)[(1+2r)(Q_j-1)+r(Q_j-1)^2]\lambda^2 - w[(1+2r)(Q_j-1)+r(1+2r)(Q_j^2-1)]\lambda \\ & + w[2r(Q_j-1)+2r^2(Q_j^2-1)] = 0\end{aligned}$$

---

[16] Any transformation of the simpler form $Q'_j = \lambda Q_j$ is ruled out by the normalization that $Q_1=1$.

The key point is that the solution for λ depends on Q_j. So λ cannot be common across choices j. Thus, it is not generally possible to find a transformation that leaves both choice probabilities of fully informed households and attitudes toward risk (and, hence, behavior of households with incomplete information) unchanged.[17] Hence, we can identify Φ_0, provided $r \neq 0$ and $\sigma_{Q_0} > 0$.

Finally, we consider some additional parameters: The relative noise in the four signals, σ_ξ, σ_τ, σ_η and σ_μ, is determined by the extent to which households update choice probabilities after receiving each type of signal. The parameters σ_ω and σ_θ are identified simply from the observed variability of prices and advertising intensities over time. In fact, these two parameters could be estimated separately from the rest of the model using just the price and advertising data. The parameters ϕ and δ govern the price-quality and advertising intensity-quality relationships. These are pinned down by the relationship between observed prices and ad intensities and the Q_j, which, as we indicated, are identified from the behavior of the well-informed households.

Note that, if households do not use price or ad intensity as signals of quality, then the inexperienced will tend to buy the more inexpensive, lower quality brands. Then, as they gather information, they will shift towards relatively higher quality brands. Conversely, if households do use price and ad intensity as signals, this pattern may be mitigated or reversed. Inexperienced households will be more likely to buy relatively expensive and/or heavily advertised brands.

III. Data

We estimate the model on Nielsen scanner panel data for ketchup, from the website of the Department of Marketing at University of Chicago. The data set records all store visits for a panel of over 3000 households in Sioux Falls, SD, and Springfield, MO, over a 153 weeks from 1986 to 1988. For each visit the data record purchases made and the price paid. TV ad exposures are available for about 60% of the households in the last 51 weeks. This is the calibration period.

We analyze the three leading brands, Heinz, Hunt and Del Monte, which together have an 84.8% market share. We ignore purchase occasions when households bought other brands. We focus on regular ketchup users by excluding households that made less than 4 purchases during the 51 weeks. We randomly select 250 households for calibration and 100 for validation. As the

[17] Note that if r=0 then (33) reduces to simply $\lambda = w/(w+\Delta)$. Then, the model is not identified without additional normalization. The normalization Φ_0=0 would be sufficient in that case, since that rules out the shift in k_1 induced by setting $w' = w + \Delta$.

sample covers 51 weeks, the calibration and holdout samples have 12750 and 5100 observations, respectively. In the calibration sample, the mean number of ketchup purchases is 8.93. The sample means of age, family size and household income are 46, 3.6 and $24,375, respectively.

As noted earlier, our model abstracts from quantity choice. Thus, we always use 32oz prices, both for the purchased brand and alternative brands, regardless of the size a household actually bought.[18] That is, we assume households compare the 32oz prices when choosing among brands. Note that 32oz is clearly the dominant size in the ketchup category. Table 1 reports the descriptive statistics for the calibration sample. Note that Heinz has the highest mean price, and the highest ad frequency (18% of households see a Heinz ad in a typical week).

The ketchup category is well suited to our purposes for several reasons. First, the literature on the price-quality relationship suggests it is stronger in frequently purchased product categories (see Rao and Monroe (1989)). Second, ketchup is a category where the brand with the high-quality positioning, that gets the highest ranking in *Consumer Reports* (1983), namely Heinz, is also the brand that has the highest mean price and highest advertising intensity. Thus, there is scope for consumers to use price and ad expenditures as signals of quality in this market. In contrast, Erdem, Keane and Sun (2004) note that advertising intensity and price are not consistently positively correlated in other categories for which scanner data is available.

IV. Empirical Results

IV.1. Model Fit and Model Selection

Our model allows for heterogeneity in the price coefficient (α_k), utility weight on quality (w_k), and risk coefficient (r_k). So we must first choose the number of types K. We estimated models with 1, 2 and 3 types, and report measures of model fit in Table 2. When we increase the number of types from one to two, the AIC and BIC improve by 95 and 80 points, respectively, and the holdout sample log-likelihood improves by 41 points. However, when we increase the number of types from two to three, and the information criteria are ambiguous (AIC improves slightly while BIC deteriorates), and the holdout log-likelihood barely improves.[19] Table 3 reports brand switching matrices for the alternative models. The homogeneous model understates

[18] Constructing prices for brands a household did not purchase is complicated, since these are not necessarily reported in the Nielsen data. We use the imputation procedure described in Erdem, Keane and Sun (1999).

[19] A notable aspect of Table 2 is that a 2-type model with myopic consumers gives a log-likelihood 63 points worse than the 2-type model with forward looking consumers. Thus, the forward-looking aspect of the model discussed in Section II.7 (i.e., consumers making trial purchases of unfamiliar brands to gather information) is important.

persistence in the data (as measured by the diagonal elements), but the two and three type models both provide an excellent fit to the switching matrix. Table 4 compares choice frequencies for the data vs. the model, and these also suggest that the two-type model fits the data well. Based on these results, we decided to use the two-type model for further analysis.

One very interesting result is already apparent in Table 3: Our model is able to capture the observed persistence in household brand choices without allowing for heterogeneity in consumer's intrinsic preferences for each brand (i.e., we do not need heterogeneous brand intercepts). Our heterogeneity is at a more fundamental level (i.e., taste for quality, risk aversion, marginal utility of consumption of the outside good). The model generates persistence in brand choices through these factors, combined with the learning process, which causes households' perceptions of brands to diverge over time as they see different signals.

IV.2. Parameter Estimates

We report parameter estimates for our preferred (two-type) model in Table 5. The price coefficient, utility weight on quality and risk parameter all have expected signs, with the latter implying that consumers are risk averse with respect to quality variation. Households in Segment 2, which is the larger segment (62%) are more price sensitive, place slightly less weight on quality, and are less risk averse with respect to quality variation than consumers in Segment 1.

The estimates of k(0) and k(1) imply that the average, across households and brands, of the "initial" perception error variance (after the 102 week initialization period), is $\sigma^2_{Q_{ij0}}$ = 0.146, suggesting that there is quality uncertainty. Regarding true quality, our estimates imply that Heinz is the highest quality, while Hunts and Del Monte are very similar. As expected, the estimates imply use experience provides more accurate signals of quality than ad exposures.

Our estimates of the slope coefficient in the price-quality and ad frequency-quality equilibrium relationships are positive (ϕ =.398, δ=.284), suggesting there is scope to use price and ad frequency as signals of quality in this market. Households perceive much greater noise in the ad frequency-quality than in the price-quality relationship (σ_μ =.532 vs. σ_η =.281), and the variability of ad frequency over time is greater than the variability of prices. These factors make price much more effective as a signal of quality than ad frequency.

Recall from our discussion of identification (section II.10) that the extent to which choice behavior differs between households with complete vs. very limited quality information is crucial for identifying r and σ_{Q_0}. In Table 6 we report simulated choice frequencies for households who know brand quality levels exactly vs. households who have not yet received any quality signals. As we would expect, greater uncertainty lowers the market share for Heinz (the higher quality brand), and raises the frequency of no purchase (since expected utility conditional on purchase is lower for a consumer with less information).

IV.3. The Role of Consumer Heterogeneity

Table 4 sheds light on the extent of the persistence in choice behavior generated by consumer heterogeneity. Note that type 1 households prefer Heinz (relatively speaking) more than type 2 households. Still, comparing Tables 3 and 4, we see that the unconditional choice frequencies for both segments are well below the diagonal elements in the switching matrix. For example, the unconditional probability of buying Hunts is 14% in segment 1, and 19% in segment 2, but the Hunts to Hunts transition rate is 55% in the full model. Thus, consumer heterogeneity cannot explain much of the high degree of persistence in brand choice generated by the model. As we will see below, most persistence is generated by the learning mechanisms.

IV.4. The Roles of the Different Information Channels

The key feature that distinguishes our model from prior work on learning is that we model four key channels through which consumers may learn about quality. How important is each channel? One way to address this question is to ask what happens to the model fit, and the behavior predicted by the model, if each channel is dropped individually from the model, or if sets of channels are dropped in tandem. We report this type of information in Table 7.

In panel A of Table 7 we report measures of model fit for various nested models that drop particular information channels. Model fit deteriorates most when we drop use experience as a signal of quality, followed by price as a signal of quality, and then by ad frequency as a signal. The smallest deterioration occurs when we drop ad content as a signal of quality.

Panel B of Table 7 provides information on how each signaling mechanism contributes to the persistence in brand choice behavior generated by the model. When we eliminate price signaling, ad frequency signaling or ad content signaling, the deterioration in the persistence generated in the model is modest. On the other hand, dropping use experience as a signal of quality leads to a substantial drop in persistence (e.g., the Hunts to Hunts transition rate drops from 54.8% in the full model to only 26.0% in model without use experience signals). Clearly, use experience signals are the main factor generating persistence in the model.

V. Policy Experiments

A key issue in marketing is whether frequent price promotion dilutes brand equity. In our model, the <u>mean</u> offer price of a brand is a signal of its quality. Frequent price promotion reduces the perceived mean price, thus reducing perceived quality. Frequent promotion also raises the variance of prices, making price a less accurate signal of quality, and increasing the perceived quality risk associated with a brand. Both factors reduce willingness to pay for a brand.

In this section, we conduct experiments that shed light on these issues. First, in Table 8, we simulate a temporary 10% price cut by Heinz lasting for one week.[20] In the week of the promotion Heinz sales increase 33%.[21] Total category sales increase by 18%, with Hunts and Del Monte sales falling by about 13% each. Thus, 80% of the short run increase in Heinz sales is due to category expansion, while only 20% is due to brand switching.

In weeks 2 through 10, when Heinz prices return to their baseline levels, its sales fall relative to the baseline, while sales of the competing brands rise. Our model has no inventory mechanism to generate this "post-promotion dip." Rather, the promotion of Heinz in week one causes consumers to revise downward their perception of the mean price of Heinz (relative to their baseline perception). This, in turn, causes consumers to revise downward their perceived quality level for Heinz, *reducing* Heinz sales for several weeks after the promotion.[22]

On the other hand, increased sales of Heinz in week 1 mean its perceived risk is lowered in week 2 (amongst those who switched to Heinz in week 1). In addition, Heinz is relatively high quality, so the extra consumers who buy it in week 1 tend to perceive it as higher quality by week 2. Both these factors tend to *increase* Heinz sales in week 2 onward. Our simulations imply that the perceived quality reducing effect of the price cut dominates the reduced risk and positive use experience effects, so that a post-promotion dip in Heinz sales does emerge.

These three mechanisms are elucidated in Table 8A. Households are divided into 7 (exhaustive) groups: (i) those who bought Heinz in period 1 under both the baseline and the promotion, (ii) those who bought Hunts under the baseline but switched to Heinz under the promotion, etc. Of the 1230 households who buy Heinz in period 1 in both cases, the number who buy Heinz in period 2 drops from 164 in the baseline to 156 under the promotion. For these

[20] The simulation is conducted in week 17. Results were very similar if we instead used week 40.
[21] This implies a short run price elasticity of demand of roughly -3.3, which is in line with previous estimates for ketchup (see Keane (1997), Erdem, Imai and Keane (2003), Erdem, Keane and Sun (2004)).
[22] Eventually, as other quality signals arrive over time, consumers' perception of Heinz quality moves back into line with their baseline perceptions. By week 10, Heinz sales return to their baseline level.

households, the only change is that the promotion lowered their perceived quality for Heinz, illustrating the 1st mechanism. On the other hand, there are three groups of consumers who switch to Heinz due to the promotion in period 1. For each of these groups, Heinz sales are *higher* in period two under the promotion than the baseline, due to the 2nd and 3rd mechanisms.

Now, returning to Table 8, we can gauge the overall importance of the three learning mechanisms. Over the whole 10-week period, Heinz sales increase 2.33%. If sales in weeks 2 through 10 had not fallen at all, the increase would have been 3.19%. Thus, 2.33/3.19 = 73% of the increased period 1 Heinz sales induced by the promotion are "incremental," while 27% represent cannibalization of future sales, due to reduced perceived quality.

It is interesting to contrast this with what Erdem, Imai and Keane (2003) obtain by fitting an inventory model to ketchup data. Their simulation (see their Table 10) implies that 20% of the temporary sales increase induced by a Heinz price promotion represents cannibalization of future sales. Thus, the inventory model implies a slightly weaker post promotion dip than does our signaling model (-20% vs. -27%). Also, in the inventory model, sales of all brands fall in period 2, not just Heinz. The other key difference is that our model implies cross-price elasticities of demand of about 1.3, while the inventory model generates cross-elasticities of only about 0.4.

In Table 8B, we simulate a temporary 10% price cut by Del Monte, the low priced brand. In the week of the promotion Del Monte sales increase 41%, implying, as expected, a larger price elasticicty of demand for the economy brand than for the premium brand, Heinz (-3.3 vs. -4.1). Note that Hunts and Heinz sales fall 5% and 2%, respectively. So, the cross-price elasticity for Heinz with respect to Del Monte is only 0.2%, while that in the opposite direction (see Table 8) is 1.3 – consistent with Blattberg and Wisniewski (1989)'s "asymmetric switching effect."

Next, we consider fundamental changes in pricing and advertising policy. Such changes generally alter consumer's purchase decision rules (see Keane (1997)). But our structural model predicts how consumers tailor their decision rules to a new environment (see Marschak (1952)).

We report our policy simulations in Table 9. In panel A, we cut Heinz mean offer price by 10% on a permanent basis. This generates a 26% increase in Heinz sales, implying an elasticity of demand with respect to permanent price changes of roughly 2.6.[23] Because price signals quality, the permanent price cut reduces perceived Heinz quality. To what extent does this detract from Heinz sales? Panel A also reports the same experiment, but holding perceived

[23] Of the incremental sales, roughly 85% is due to category expansion, while 15% is due to brand switching.

Heinz quality fixed at baseline levels. Then, the increase in sales is 32%. Thus, reduced quality perceptions counteract 6 points (or about a fifth) of the increase in Heinz sales that would have occurred if price could have been cut without altering quality perceptions. This suggests that price plays an important role in signaling quality.[24]

In Table 9, Panel B, we decrease Heinz price variability by 20% while holding its mean offer price fixed.[25] Lower price variability makes price a more accurate signal of quality, which may enhance perceived quality for Heinz (since it is positioned as high quality), while also reducing perceived risk. But these factors are dominated by the direct effect of less frequent price promotion, so Heinz total sales over the sample period fall by 7.9%.

In Panel C of Table 9 we implement experiments that mimic a switch to an "every day low pricing" (EDLP) strategy, i.e., simultaneous mean offer price and price variability reductions that leave total Heinz sales unchanged. As total sales are unchanged, it is reasonable to assume that cost of goods sold is roughly fixed in these experiments.[26] Then, increases in mean accepted price generate increases in profits. According to our simulations, Heinz could reduce mean price by 2% while reducing price variability 48%, and this would increase revenues by 0.47%. If the original price/cost margin were, say, 20%, this would induce a 2.5% increase in profits.

This experiment is suggestive that an EDLP policy can enhance brand equity, and it is interesting that such policies were widely adopted by many retailers shortly after our sample period. However, it should be stressed that our partial equilibrium model does not predict the impact of possible competitor reactions to such a change in Heinz pricing policy.

Finally, in Table 9, Panel D, we simulate a 50% increase in ad intensity by Heinz. This enhances perceived quality for Heinz both because (i) high ad frequency signals quality, and (ii) households now receive more frequent ad content signals of Heinz' high quality position. As a result, Heinz sales are predicted to rise by 17%. Of course, for a low quality brand, effects (i) and (ii) would work against each other instead of being reinforcing.

---

[24] Indeed, we calculate that the average perceived quality of Heinz falls from .504 under the baseline to .457 with the permanent price reduction. This means that roughly 40% of its perceived quality advantage over Hunts is dissipated.

[25] We do this as follows: 1) Find mean offer price for each brand. 2) Scale up the deviations of price realizations from that mean so as to achieve the desired increase in variance. 3) Determine how this transformation affects mean and variance of the log price. 4) Modify the log price equation accordingly so as to keep the mean price fixed in levels. 5) Simulate behavior given the new price data and the new price process.

[26] Of course, it is now well understood that reduced variability in prices, and hence in sales, leads to less inventory along the supply chain, reducing inventory costs (see, e.g., Ohno (1988)). Hence, costs may fall despite fixed overall sales if price variability is reduced. Indeed, such supply chain considerations were presumably the main reason for EDLP adoption.

VI. Discussion and Conclusions

We have proposed and estimated a dynamic brand choice model in which consumers learn about brand quality through four distinct channels: 1) price signaling quality, 2) advertising frequency signaling quality, 3) use experience providing direct (but noisy) information about quality, 4) advertising providing direct (but noisy) information about quality. The model was estimated on Nielsen scanner data for the ketchup category and it appears to fit the data well.

Our estimates imply that mean offer price plays a very important role in signaling brand quality.[27] This implies that frequent price promotions, that reduce the perceived mean offer price of a brand, can feed back and adversely impact perceived quality. Simulations of the model imply that roughly one quarter of the increase in sales generated by a temporary price cut represents cannibalization of future sales due to the brand equity diluting effect of the promotion.

The post-promotion dip generated by the price/quality signaling mechanism in our model looks very similar to that generated by an inventory model (see Erdem, Imai and Keane (2003)). Future work is needed to help distinguish between these two ways of interpreting the data.

Our findings also suggest that reductions in mean offer price combined with reductions in price variability, as in an EDLP policy, can potentially lead to increased profitability. But, since our partial equilibrium model does not incorporate competitor reaction to changes in pricing policy, this result is only suggestive and should be interpreted with caution.

The most frequent criticism of learning models like Erdem and Keane (1996) and the present paper is the *a priori* judgment that learning is "not important" for mundane categories like detergent and ketchup. This is a view with which we strenuously disagree. First, intuitively, we believe that most consumers who are very "experienced" in categories like detergent or ketchup are in fact very familiar with just one (or a small number) of brands that they buy frequently - leaving them very unfamiliar with the alternatives. This is exactly the mechanism through which learning models generate brand equity for the preferred familiar brand via the risk term. Second, learning models fit the dynamics of choice behavior in "mundane" frequently purchased categories very well. The scientific challenge for one who finds these models *a priori* implausible is to find an alternative mechanism that will explain dynamics equally well.

---

[27] Ranking signaling mechanisms by their order of importance, as measured by the deterioration in model fit when each mechanism is excluded, our results suggest that use experience is the most important signal of quality, followed by price, then ad frequency and then ad content. However, all four mechanisms appear to be important, since dropping any one of them led to a significant deterioration in model fit.

References

Abernethey, Avery M. and George R. Franke (1996), "The Information Content of Advertising: A Meta-Analysis," *Journal of Advertising*, 25 (2), 1-17.

Ackerberg, Dan (2003), "Advertising, Learning and Consumer Choice in Experience Good Markets: A Structural Empirical Examination," *International Economic Review*, 44, 1007-40.

Anand, Bharat, and Ron Shachar (2002): "Risk Aversion and Apparently Persuasive Advertising." Harvard Business School Working Paper Series, No. 02-099.

Blattberg, Robert and Kenneth Wisniewski (1989), "Price-Induced Patterns of Competition" *Marketing Science* 8 (4), 291-309.

Caves, R.E. and D.P. Greene (1996), "Brands' quality levels, prices and advertising outlays: Empirical evidence on signals and information costs," *International Journal of Industrial Organization,* 14 (1), 29-52.

Ching, Andrew (2002): "Consumer Learning and Heterogeneity: Dynamics of Demand for Prescription Drugs After Patent Expiration," Working Paper, Ohio State University.

Crawford, Gregory S. & Shum, Matthew (2003): "Uncertainty and Learning in Pharmaceutical Demand: Anti-Ulcer Drugs," *Econometrica*, forthcoming.

DeGroot, Morris H. (1970), *Optimal Statistical Decisions,* McGraw-Hill.

Erdem, Tülin and Michael P. Keane (1996), "Decision-Making under Uncertainty: Capturing Choice Dynamics in Turbulent Consumer Goods Markets", *Marketing Science*, 15, 1-21.

Erdem, Tülin, Michael Keane and Baohong Sun (2004), "The Impact of Advertising on Consumer Price Sensitivity in Experience Goods Markets," Working paper.

Erdem, Tülin, Susumu Imai and Michael Keane (2003), "A Model of Consumer Brand and Quantity Choice Dynamics under Price Uncertainty," *Quantitative Marketing and Economics*, 1 (1), 5-64.

Erdem, Tülin and Baohong Sun (2002), "An Empirical Investigation of Spillover Effects of Marketing Mix Strategy in Umbrella Branding," *Journal of Marketing Research,* 39 (4), 408-420.

Farell, J. (1980), "Prices as Signals of Quality", Ph.D. thesis, University of Oxford.

Gerstner, Eitan (1985), "Do higher Prices Signal Higher Quality?", *Journal of Marketing Research*, 22 (May), 209-215.

Hendel, Igal and Aviv Nevo (2003), "Measuring the Implications of Sales and Consumer Stockpiling," working paper, University of California, Berkeley.

Hong, Pilky, R. Preston McAfee and Ashish Nayyar. (2002) "Equilibrium Price Dispersion with Consumer Inventories," Journal of Economic Theory, forthcoming.

Jedidi, Kamal, Carl Mela and Sunil Gupta (1999), "Managing advertising and promotion for long-run profitability," *Marketing Science*, 18 (1), 1-22.

Keane, Michael P. and Kenneth I. Wolpin (1994), "Solution and Estimation of Dynamic Programming Models by Simulation", *Review of Economics and Statistics*, 76 (4), 648-672.

Keane, Michael P. (1997). "Current Issues in Discrete Choice Modeling," Marketing Letters.

Kihlstrom, Richard E. and Michael H. Riordan (1984), "Advertising as a Signal", *Journal of Political Economy*, 92, 427-450.

Lichtenstein, Donald R. and Scott Burton (1989), "The Relationship between Perceived and Objective Price-Quality", *Journal of Marketing Research,* 26, (November), 429-443.

Marschak, J. (1952). "Economic Measurements for Policy and Prediction." In W.C. Hood and T.J. Koopmans (eds.), Studies in Econometric Metho. New York: Wiley. 1-26.

Mehta, N., S. Rajiv, K. Srinivasan (2003), "Price uncertainty and consumer search: A structural model of consideration set formation," *Marketing Science*, 22 (1), 58-84.

Milgrom, Paul and John Roberts (1986), "Prices and Advertising Signals of Quality," *Journal of Political Economy, 94,* 796-821.

Ohno, Taiichi (1988). "Toyota Production System: Beyond Large Scale Production," Productivity Press, New York.

Rao, Akshay R. Kent B. Monroe (1989), "The effect of price, brand name and store name on buyers' perceptions of product quality: an integrative review, " *Journal of Marketing Research*, 26 (July), 351-357.

Resnick, Alan and Bruce L. Stern (1977), "An Analysis of Information Content in Television Advertising," *Journal of Marketing* 52 (1), 50-53.

Rust, J. (1987). "Optimal replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55(5), 999-1033.

Spence, Michael (1974), *Market Signaling*, Cambridge, MA: Harvard University Press.

Tirole, Jean (1990), *The Theory of Industrial Organization*, Cambridge, MA: MIT Press.

Winer, Russell S. (1986), "A Reference Price Model of Brand Choice for Frequently Purchased Products," *Journal of Consumer Research*, 13 (2), 250-256.

Table 1: Descriptive Statistics

Brand Name	Market Share	Mean Offered Price[1]	Mean Accepted Price[1]	Mean Weekly Advertising Frequency[2]	Mean Number of Advertising Exposures[3]
Heinz	66.15%	$1.349	$1.302	0.180	2.12
Hunts'	17.26%	$1.197	$1.141	0.096	1.57
Del Monte	16.58%	$1.184	$1.104	0	0

1. Prices are normalized at 32oz per bottle.
2. The percentage of households who see at least one ad for the brand in a typical week
3. The mean number of ads seen in a given week, conditional on ad exposure.

Table 2: Model Selection

		Myopic Model with Learning (Two types)	Dynamic Models		
			One Type	Two Types	Three Types
In-sample (Sioux Falls):	-LL	11854.0	11890.2	11791.1	11779.0
	AIC	11882.0	11914.2	11819.1	**11811.0**
	BIC	11986.3	12003.6	**11923.4**	11930.3
Out-of-sample (Springfield):	-LL	4960.1	4872.7	4832.0	**4829.2**

* Calibration sample: Number of observations = 12750. Number of households = 250. Number of periods = 51.
Holdout sample: Number of observations = 5100. Number of households = 100. Number of period = 51.

** Note: AIC=-Log-likelihood+# of parameters.
BIC=-Log-likelihood+0.5*# of parameters*ln(# of observations).

***There are 28, 24, 28, and 32 parameters in the four estimated models.

Table 3: Comparison of Simulated and Sample Frequencies

	Data	Model Prediction		
		One Type	Two Types	Three Types
Heinz	11.61% (66.15%)	10.73%(64.48%)	11.68% (66.78%)	11.98%(67.71%)
Hunts	3.03% (17.26%)	3.09%(18.57%)	2.94% (16.81%)	3.16%(17.70%)
Del Monte	2.91% (16.58%)	2.82%(16.95%)	2.87% (16.41%)	2.71%(15.18%)
No Purchase	82.45%	83.36%	82.51%	82.15%
Segment 1			38.4%	28.0%
Heinz			13.22% (71.77%)	14.35%(75.57%)
Hunts			2.61% (14.17%)	2.50%(13.16%)
Del Monte			2.59% (14.06%)	2.14%(11.27%)
No Purchase			81.60%	81.01%
Segment 2			61.6%	42.0%
Heinz			10.73% (63.76%)	12.61%(68.76%)
Hunts			3.15% (18.60%)	3.02%(16.47%)
Del Monte			3.04% (18.00%)	2.71%(14.78%)
No Purchase			83.08%	81.66%
Segment 3				30.0%
Heinz				8.89%(55.19%)
Hunts				3.97%(24.67%)
Del Monte				3.24%(20.14%
No Purchase				83.90%

Note: Probabilities conditional on purchase are in parenthesis.
In the two segment model, the segment proportions are 38.4% and 61.6%. In the three segment model, the segment proportions are 28.0%, 42.0% and 30.0%.

Table 4: Brand Switching Matrices

Sample	Myopic Model With 2 Types	One Type Model	Two Type Model	Three Type Model
0.901 0.072 0.027 0.302 0.530 0.168 0.337 0.260 0.403	0.891 0.073 0.036 0.295 0.535 0.170 0.351 0.273 0.376	0.891 0.061 0.048 0.321 0.482 0.197 0.354 0.293 0.353	0.907 0.068 0.025 0.294 0.548 0.192 0.314 0.285 0.401 Segment 1 (38.4%) 0.918 0.063 0.019 0.289 0.554 0.157 0.306 0.277 0.417 Segment 2 (61.6%) 0.900 0.071 0.029 0.241 0.533 0.213 0.319 0.290 0.391	0.908 0.061 0.031 0.302 0.527 0.171 0.347 0.253 0.400 Segment 1 (28.0%) 0.919 0.059 0.020 0.285 0.550 0.150 0.309 0.280 0.423 Segment 2 (42.0%) 0.897 0.067 0.026 0.250 0.504 0.222 0.322 0.286 0.379 Segment 3 (30.0%) 0.906 0.054 0.048 0.391 0.548 0.119 0.417 0.182 0.408

Table 5: Structural Model Estimation Results

Parameters that Differ by Consumer Segment				
	Segment 1		Segment 2	
	Estimate	Std. Error	Estimate	Std Error
Price coefficient (α)	-0.111	0.06	-0.307	0.06
Utility weight (w)	1.992	0.33	1.606	0.41
Risk coefficient (r)	-0.363	0.09	-0.247	0.10
Segment membership probability (π)	0.384	0.10	0.616	----

Homogenous Parameters		
	Estimate	Std. Error
k(0)	-0.661	0.29
k(1)	0.066	0.02
Quality Levels:		
Q_{Heinz}	0.501	----
Q_{Hunts}	0.393	0.12
$Q_{Del\ Monte}$	0.368	0.08
Use experience signal variability (σ_ξ)	0.292	0.12
Advertising message variability (σ_τ)	0.612	0.19
Price Signaling Quality Equation:		
Intercept (P_0)	-0.031	0.01
Slope (ϕ)	0.398	0.12
Brand Specific Constants:		
η_{Heinz}	0.053	0.02
η_{Hunt}	-0.025	0.01
Standard deviation of η across brands (σ_η)	0.281	0.09
Price variability (σ_w)	0.401	0.13
Advertising Signaling Quality Equation:		
Intercept (A_0)	-0.194	0.27
Slope (δ)	0.284	0.10
Brand Specific Constants:		
μ_{Heinz}	0.085	0.023
μ_{Hunt}	-0.005	0.003
Standard deviation of μ across brands (σ_μ)	0.532	0.19
Box-Cox parameter (β)	0.621	0.20
Advertising (frequency) variability (σ_θ)	0.532	0.15
Utility from No Purchase Option:		
Intercept (Φ_0)	0.983	0.10
Time trend (Φ_1)	0.001	0.13

Table 6: Comparison of Choice Probabilities for Households with Different Degrees of Knowledge about Quality

Perception Error Variance	Average Purchase Probabilities			
	Heinz	Hunt	Del Monte	No Purchase
$\sigma^2_{Q_0} = 0.146$	11.68%	2.94%	2.87%	82.51%
$\sigma^2_{Q_0} = 0$	13.62%	2.58%	2.52%	81.28%

Note: 0.146 is the average initial perception error variance, across consumers and brands. The estimate of k(0) implies that the prior standard deviation before any learning takes place is exp(-.661) = 0.5163, giving a prior variance of 0.267. Use experience in our 102 week initialization period reduces this, on average, to 0.146.

Table 7: Comparing the Importance of the Information Channels

A. Model Fit

		No Price Signaling Quality	No Ad Frequency Signaling	No Ad Content Signaling	No Use Experience Signaling	Price as Only Signal of Quality	Full Model
In-sample (Sioux Falls):	-LL	11950.5	11929.1	11904.7	11996.8	12051.2	11791.1
	AIC	11973.5	11952.1	11931.7	12003.8	12071.2	11819.1
	BIC	12059.2	12037.8	12032.3	12124.4	12145.7	11923.4
Out-of-sample (Springfield):	-LL	4902.0	4881.8	4879.9	4922.8	4940.1	4832.0

B. Brand Switching Matrices

No Price Signaling Quality	No Ad Frequency Signaling	No Ad Content Signaling	No Use Experience Signaling	Price as Only Signal of Quality	Full Model
0.857 0.074 0.069	0.864 0.081 0.055	0.883 0.062 0.055	0.720 0.214 0.066	0.601 0.295 0.104	0.907 0.068 0.025
0.280 0.462 0.258	0.318 0.453 0.229	0.320 0.481 0.199	0.444 0.260 0.296	0.427 0.192 0.381	0.294 0.548 0.192
0.349 0.289 0.362	0.359 0.302 0.339	0.368 0.288 0.344	0.383 0.378 0.239	0.441 0.378 0.181	0.314 0.285 0.401

Table 8: Effects of Temporary 10% Heinz Price Decrease in Week 17

Week	Change of Average Purchase Probabilities			
	Heinz	**Hunt**	**Del Monte**	**Total**
17	**32.91**	**-12.90**	**-12.70**	**17.65**
18	**-3.50**	**2.35**	**1.99**	**-1.61**
19	**-2.23**	**1.21**	**0.94**	**-1.17**
20	**-1.33**	**0.85**	**0.32**	**-0.79**
21	**-0.80**	**0.42**	**0.13**	**-0.44**
22	**-0.54**	**0.28**	**0.03**	**-0.32**
23	**-0.30**	**0.10**	**0.02**	**-0.19**
24	**-0.15**	**0.04**	**0.01**	**-0.09**
25	**-0.04**	**0.01**	**0.00**	**-0.02**
26	**-0.01**	**0.00**	**0.00**	**-0.01**
Cumulative over 10 weeks	**2.33**	**-0.85**	**-0.95**	**1.29**
Cumulative, assuming no change after week 17	**3.19**	**-1.35**	**-1.28**	**1.75**

Note: The table reports the percentage change of average probabilities for each brand by week, following a temporary 10% price cut for Heinz in week 17, compared to a baseline simulation under the present pricing policy. The average probabilities are calculated using a sample of 10,000 hypothetical consumer histories simulated from the model. "Week 1" of the simulation is actually week 17 in the data. We choose week 17 as the base period for the simulation because all brands were selling at roughly their average prices during that week (i.e., there were no sales in the baseline).

Table 8A. Breakdown of Effect of Sale in Table 8

Choice in Period 17			Choice in Period 18					
Baseline	**Simulation**	**Number of People**	**Baseline**			**Simulation**		
			Heinz	Hunts	Del Monte	Heinz	Hunts	Del Monte
Heinz	**Heinz**	1.230	164	17	10	156	14	8
Hunts	**Heinz**	43	8	3	0	10	1	0
Hunts	**Hunts**	287	7	10	5	3	5	2
Del Monte	**Heinz**	36	4	2	2	5	0	1
Del Monte	**Del Monte**	246	6	5	11	4	3	12
No Purchase	**Heinz**	325	56	6	4	61	4	2
No Purchase	**No Purchase**	7,833	925	297	208	890	321	218
Total		10,000	1,170	340	240	1,129	348	245

Note: We report the counts of consumers in each cell, based on a simulation of 10,000 hypothetical consumers. In period 17, the choice frequencies are Heinz 12.30%, Hunts 3.30%, Del Monte 2.80% under the baseline, and Heinz 16.34%, Hunts 2.87% Del Monte 2.46% under the simulation. Note that the changes in choice frequencies implied by these figures differ slightly from those in Table 8. This is because the figures in Table 8 are based on the choice probabilities implied by the model, averaged across 10,000 simulated consumers.

Table 8B: Effects of Temporary 10% Del Monte Price Decrease

Week	Change of Average Purchase Probabilities			
	Heinz	**Hunt**	**Del Monte**	**Total**
17	**-2.21**	**-5.01**	**41.11**	**3.96**
18	**0.21**	**1.13**	**-3.53**	**-0.12**
19	**0.11**	**1.01**	**-2.21**	**-0.17**
20	**0.11**	**0.74**	**-1.94**	**-0.11**
21	**0.07**	**0.49**	**-1.18**	**-0.10**
22	**0.03**	**0.23**	**-0.59**	**-0.05**
23	**0.01**	**0.02**	**-0.19**	**-0.02**
24	**0**	**0.01**	**-0.03**	**-0.00**
25	**0**	**0**	**0.01**	**0.00**
26	**0**	**0**	**0**	**0**
Cumulative over 10 weeks	**-0.17**	**-0.19**	**3.18**	**0.34**
Cumulative, assuming no change after week 17	**-0.21**	**-0.53**	**4.15**	**0.39**

Table 9: Effects of Permanent Changes in Heinz Pricing Policy on Sales

A. Cut Heinz's price by 10% on a permanent basis.

	Heinz	Hunt	Del Monte	Total
Change of Purchase Probability	**26.42**	**-8.37**	**-7.26**	**14.85**
Change of Purchase Probability without Changing Perceived Quality	**32.23**	**-12.16**	**-11.90**	**27.04**

B. Decrease Heinz's price variability by 20% while holding mean price fixed.

	Heinz	Hunts	Del Monte	Total
Change of Purchase Probability	**-7.87**	**4.23**	**5.29**	**-4.05**

C. Combine cut in mean Heinz offer price with decrease in price variability to leave sales unchanged

Percentage Cuts in Mean Offer Price of Heinz	Decrease in Price Variability	Percentage Change of Average Accepted Price
-2%	**48%**	**0.47%**
-4%	**81%**	**0.68%**

D. Increase Heinz's advertising intensity by 50%

	Heinz	Hunt	Del Monte	Total
Change of Purchase Probability	**17.24**	**-6.74**	**-4.23**	**9.52**

The table reports the percentage changes in each of the indicated quantities for the period after the policy change, compared to a baseline simulation under the present pricing policy.