BAOHONG SUN, SCOTT A. NESLIN, and KANNAN SRINIVASAN*

Logit choice models have been used extensively to study promotion response. This article examines whether brand-switching elasticities derived from these models are overestimated as a result of rational consumer adjustment of purchase timing to coincide with promotion schedules and whether a dynamic structural model can address this bias. Using simulated data, the authors first show that if the structural model is correct, brand-switching elasticities are overestimated by stand-alone logit models. A nested logit model improves the estimates, but not completely. Second, the authors estimate the models on real data. The results indicate that the structural model fits better and produces sensible coefficient estimates. The authors then observe the same pattern in switching elasticities as they do in the simulation. Third, the authors predict sales assuming a 50% increase in promotion frequency. The reduced-form models predict much higher sales levels than does the dynamic structural model. The authors conclude that reduced-form model estimates of brand-switching elasticities can be overstated and that a dynamic structural model is best for addressing the problem. Reduced-form models that include incidence can partially, though not completely, address the issue. The authors discuss the implications for researchers and managers.

Measuring the Impact of Promotions on Brand Switching When Consumers Are Forward Looking

Logit choice models have proved an invaluable tool for studying consumer decision processes. The models have generated insights into issues such as market segmentation (Bucklin, Gupta, and Siddarth 1998; Chintagunta, Jain, and Vilcassim 1991; Kamakura and Russell 1989), state dependence (Guadagni and Little 1983; Keane 1997a, b; Seetharaman, Ainslie, and Chintagunta 1999), and competitive asymmetries (Allenby and Rossi 1991). A fundamental finding is that promotions influence consumer choice; that is, they cause consumers to switch from Brand A to Brand B. Gupta (1988), Chintagunta (1993), Chiang (1995), Bucklin, Gupta, and Siddarth (1998), and Bell, Chiang, and Padmanabhan (1999) have examined the magnitude of the brand-switching effect, at least relative to dynamic effects such as stockpiling, and have found that brand switching accounts for the majority of the current period promotion effect.

A related research stream has studied consumers' dynamically rational decisions about quantity and purchase timing: Golabi (1985), Assunção and Meyer (1993), Meyer and Assunção (1990), Helsen and Schmittlein (1992), and Krishna (1992, 1994a, b) study how promotion uncertainty determines consumers' optimal forward-buying behavior. Erdem and Keane (1996) and Gönül and Srinivasan (1996) develop stochastic structural dynamic programming models to accommodate consumers' forward-looking behavior. Gönül and Srinivasan find that consumers can accelerate or decelerate purchases to coincide with a promotion schedule as long as they have sufficient inventory. Erdem, Imai, and Keane (2003) develop a dynamic structural model in which consumers form future price expectations and decide when, what, and how much to buy. These articles demonstrate that dynamic structural models are important for capturing

*Baohong Sun is Assistant Professor of Marketing, Kenan-Flagler Business School, University of North Carolina (e-mail: sunb@bschool.unc.edu). Scott A. Neslin is Albert Wesley Frey Professor of Marketing, Tuck School of Business, Dartmouth College (e-mail: Scott.Neslin@Dartmouth.edu). Kannan Srinivasan is H.J. Heinz II Professor of Management, Marketing, and Information Systems, Graduate School of Industrial Administration, Carnegie Mellon University (e-mail: kannans@andrew.cmu.edu). The authors thank Tülin Erdem, Eric Bradlow, and participants in the marketing seminar at University of Michigan; the Tuck School Marketing Workshop; National University of Singapore; University of California, Riverside; Harvard University; Yale University; Washington University; and Duke University and in the Marketing Summer Camp at Carnegie Mellon University for valuable comments on previous versions of this article. They also thank Andrew Hayes, Tuck School of Business, for valuable research assistance. The authors thank the three anonymous *JMR* reviewers for several constructive suggestions.

Journal of Marketing Research
Vol. XL (November 2003), 389–405

rational consumer adjustments in purchase timing due to expectations of future promotion activity.

Our research builds on the study of the brand-switching effect by (1) demonstrating a potential upward bias in logit-estimated switching elasticities and (2) showing that a dynamic rational model can significantly alleviate this bias. The intuition behind the bias is that the logit choice model does not take into account consumers' rational purchase-timing adjustments made to take advantage of a deal. The bias misidentifies the promotion purchases as brand switches and therefore overestimates the switching effect. Keane (1997a, p. 311) describes the situation as follows: "Suppose after a few years of shopping I realize that my favorite supermarket always cuts the price of my favorite detergent, say Era, from $4.00 to $3.50 during one week of every month. Most other shoppers at the store realize this too. So those who like Era adopt a decision rule that says: 'Only buy Era if the price is $3.50, and then buy enough so that my inventory will last at least 7 weeks (the longest possible time until the next sale)'.... If one were to estimate a MNL [multinomial logit] model for detergent choice in the store, the price elasticity of demand for Era would appear enormous."

The subsequent numerical example is motivated by Keane's (1997a) research.[1] Consider two consumers: One prefers Brand A and one prefers Brand B. In a no-promotion environment, their purchases might be as follows:

Week	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
Consumer 1	A		A		A		A	
Consumer 2	B		B		B		B	

Now consider that Brand A is promoted in Weeks 2 and 6. Assume that Consumer 2 does not react to this promotion; there is no switching effect. Consumer 1 reacts by accelerating the second purchase forward from Week 3 to Week 2, not purchasing in Week 5 in anticipation of a promotion, and purchasing in Week 6 when the promotion occurs:

Week	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
Consumer 1	A	A				A	A	
Consumer 2	B		B		B		B	

Note that the promotion has not changed consumers' brand choices; it has only changed the timing of Consumer 1's purchase. However, when a logit model is used, it reads data from purchase observations only and associates higher probabilities of purchasing Brand A with promotions available for Brand A. The model overestimates the brand-switching effect because it is designed to "fit" the statistical relationship between purchase probabilities and promotion availability without recognizing the purchase-time adjustment of a forward-looking consumer.[2]

A nested logit model may partly alleviate the overestimation problem by calculating brand switching conditional on a promotion-driven purchase-timing decision. However, this model is a reduced form in that it does not model the process of dynamic decision making: It still "fits" the relationship between probability of purchasing and promotion availabilities of Brand A. In general, logit and nested logit choice models are too simple to take into account consumers' forward-looking behavior. For example, nested logit models naturally capture purchase acceleration but do not easily capture purchase deceleration. A dynamic structural model in which the purchase-timing decision is endogenized is designed to model this behavior explicitly.

In addition to more accurate estimation of brand switching as a result of marginal additional promotion, another potential advantage of a dynamic rational model is that it can avoid overprediction of brand switching when there is a shift in promotion policy. Keane (1997a) conjectures that the logit model would be especially ineffective if a brand were to change its promotion policy, because the model does not describe how consumer purchase strategies adapt to the new policy.

The intuition that logit choice models overstate the switching effect makes sense but simplifies the problem in two ways. First, logit models completely ignore the purchase-timing decision and consider promotion solely as a brand-switching game. As we mentioned previously, techniques exist to examine choice and timing decisions jointly (Bell, Chiang, and Padmanabhan 1999; Chiang 1991; Chintagunta 1993; Gupta 1988). Second, the intuition does not explicitly take into account consumer preferences. Modeling heterogeneity is an inherent part of any logit model and may help alleviate the bias.

Therefore, we undertake the following steps to demonstrate the existence of bias and prescribe a solution for it: First, we develop a dynamic structural consumer-decision model of incidence and choice. The potential advantage of the structural model is more accurate brand-switching elasticity estimates because it endogenously models the process by which rational consumers adjust their timing decisions in response to promotion uncertainty. In accordance with Keane (1997a), we believe this will be particularly effective in evaluating changes in promotion policy. Second, we compare this model to variations of both stand-alone logit choice models and choice/incidence models, taking into account heterogeneity in consumer preferences. We conduct the following analyses:

- Synthetic data simulation: Assuming that the structural model is correct, we generate purchase-timing and choice decisions, and we estimate logit, nested logit, and structural models on the data. This simulation shows that the logit and nested logit models produce upwardly biased estimates of promotion-induced brand switching, compared with the true effect as generated by the structural model. The logit model results are more upwardly biased than those of the nested logit model, suggesting that taking into account purchase timing at least partially addresses the bias issue.
- Empirical estimation: We estimate the logit, nested logit, and structural models on real data and compare brand-switching elasticities across the models. We find that the structural model fits the data better, and the elasticity results mirror the synthetic data simulation. The logit model elasticities are greater than the nested logit elasticities, which in turn are greater than the structural model elasticities.
- Policy simulation: Starting with the real data, we increase the promotion frequency for one brand and use our models to predict sales with and without the policy change. This is a test of

[1]We acknowledge the stimulating comments raised by an anonymous reviewer on a different article by one of the authors, which also served to motivate this example.

[2]Note that promotions can also induce larger purchase quantities, and it is not difficult to construct examples that show that this effect can also induce an upward bias in brand-switching effects. This therefore may accentuate the bias. However, we simplify the analysis by focusing on brand choice and purchase incidence.

the Lucas critique (Lucas 1976), which states that policy models that do not account for adjustments agents make in response to a "regime" change produce biased predictions. As we expected, we find that the logit and nested logit models predict much higher sales for the brand with increased promotion than does the structural model; this is because the structural model addresses the Lucas critique by modeling how consumers change their purchase timing in response to policy change.

Overall, the results suggest that logit choice models overestimate brand-switching elasticities. Although nested logit models partially mitigate the problem, a structural model, such as the one we propose, addresses the problem comprehensively.

We proceed as follows: First, we develop the structural model. Second, we present our synthetic data simulation, our empirical estimation, and our policy simulation. Finally, we conclude with a summary and discussion of the implications for both researchers and managers.

MODELING CONSUMER DECISIONS IN A STRUCTURAL FRAMEWORK

Dynamic Structural Decision Models

There is a growing body of literature that establishes that consumers are forward looking in their purchase decisions. The evidence includes the observation that consumers stockpile product in response to promotions (Blattberg, Eppen, and Lieberman 1981; Neslin, Henderson, and Quelch 1985). The apparent reason consumers stockpile is that they realize the lower price, which is available only temporarily, is worth the additional future inventory storage costs. Several articles have models that formalize this process (Assunção and Meyer 1993; Blattberg, Eppen, and Lieberman 1981; Erdem, Imai, and Keane 2003; Gönül and Srinivasan 1996; Krishna 1992, 1994a, b; Sun 2003).

We propose a dynamic structural model for analyzing the effect of promotion on brand switching in an uncertain promotion environment. Our primary theme is that forward-looking, rational consumers make it difficult for nondynamic models to measure brand switching. Therefore, the key phenomena we incorporate are promotion expectations, brand choice and incidence, and dynamic utility maximization. The need to model promotion expectations and brand choice is obvious given our purpose. The dynamic rational model literature assumes that consumers either maximize long-term utility or minimize long-term costs under a weekly consumption constraint. We believe that utility maximization is more appropriate for our needs, because we want consumers to have the option of forgoing or accelerating consumption.

Comparing these aspects with those in previous literature, Blattberg, Eppen, and Lieberman (1981) focus on promotion but assume that the promotion schedule is known; in addition, they do not model brand choice, and they assume that the consumer minimizes costs subject to a consumption constraint. Assunção and Meyer (1993) maximize consumer utility and model promotion expectations but do not model brand choice. Krishna (1992, 1994a, b) considers brand choice and models promotion expectations but assumes that customers minimize costs subject to a consumption constraint, as do Blattberg, Eppen, and Lieberman. Erdem and Keane (1996) maximize utility and examine brand choice but consider product attributes, not promotion. Gönül and Srinivasan (1996) consider promotion expectations, but do not model brand choice, and consider cost minimization. Erdem, Imai, and Keane (2003) consider brand choice and purchase quantity in a dynamic utility maximization framework. Our model is similar to theirs in that it also draws on the general framework of Hanemann (1984) to model utility for consumption. However, we tailor our model to investigate the impact of promotion on brand switching. For example, we treat heterogeneity using a continuous distribution rather than the latent class approach adopted by Erdem, Imai, and Keane. We also use a simpler promotion expectations model that focuses on the promotion schedule of individual brand-sizes. This article differs from that of Sun (2003), who endogenizes the consumption decision and investigates promotion effect on endogenous consumption.

Model Specification

We model the consumer decision process as a dynamic programming problem under promotion uncertainty. We assume that consumer i = 1, ..., I visits the store on the periodic (e.g., weekly) basis t = 1, ..., T. On each store visit, the consumer decides whether to buy and which brand to buy given the observed promotions and expected future promotions of the j = 1, ..., J brand-sizes. We denote the no-purchase choice by j = 0. The objective of consumer i is to determine, for each of a series of store-visit occasions, whether to purchase the category and if so which brand, so as to maximize the sum of discounted expected future utility, U_{it}, over the finite planning horizon.

$$\max_{d_{ijt}} E\left\{\sum_{\tau=t}^{T} \delta^{\tau-1} U_{i\tau}\right\}. \tag{1}$$

The indicator variable $d_{ijt} = 1$ if consumer i chooses brand j at time t, and $d_{ijt} = 0$ otherwise for j = 1, ..., J. Because j = 0 indexes the category decision, $d_{i0t} = 1$ means that consumer i decides not to buy the product category (i.e., "chooses" alternative "0"), and $d_{i0t} = 0$ means that consumer i decides to buy the product category in week t. The length of the finite decision horizon is T, and $0 < \delta < 1$ is the discount factor to reflect that consuming now is preferred to consuming later (Erdem and Keane 1996; Gönül and Srinivasan 1996). We denote the mathematical expectation operator as $E(\cdot)$.

For each week t, we define the consumer's utility function over that week's consumption of each brand-size, C_{ijt}, and of consumption of a composite of other goods, Z_{it}. Thus, the utility is given by

$$U_{it} = \sum_{j=1}^{J} \psi_{ijt} C_{ijt} + \alpha_{1i} Z_{it}, \tag{2}$$

where α_{1i} measures the benefit from consuming the composite of other goods.[3] The term ψ_{ijt} denotes the consumption benefit associated with consuming brand j. We model this as follows:

$$\psi_{ijt} = \alpha_{2ij} + \alpha_{3i} Last_{ijt}. \tag{3}$$

[3]Note that we assume that the composite good is not storable, so we view Z_{it} as either purchase or consumption. This is to simplify our model, the purpose of which is to explain behavior in the focal category.

The utility function is state dependent and stochastic, as Hanemann (1984) introduces. The variable $Last_{ijt}$ indicates whether consumer i purchased brand j at purchase occasion t − 1. The benefit from consumption of brand j depends on $Last_{ijt}$. The coefficients α_{2ij} measure consumer i's intrinsic preference for brand j. The parameter α_{3i} captures that consumers' learning of the quality of a product or their familiarity with a product can influence their consumption preference for a product. This state dependence phenomenon is also referred to as purchase-event feedback (Ailawadi, Gedenk, and Neslin 1999).

We normalize the price of the composite goods to be one. At each period, consumer i has income y_{it} and incurs costs as a result of purchasing brands in the category, consuming other goods, managing category inventory, and incurring stockouts, in accordance with budget constraint:[4]

$$(4)\qquad y_{it} = \sum_{j=1}^{J} P_{ijt}d_{ijt} + Z_{it} + \delta_{1i}\sum_{j=1}^{J}Inv_{ijt} + \delta_{2i}I\left(\sum_{j=1}^{J}Inv_{ijt} - \bar{C}_i\right),$$

where

$$(5)\qquad P_{ijt} = Price_{ijt} - Prom_{ijt}Dscnt_{ijt}.$$

The variable P_{ijt} is the net price paid by consumer i, $Price_{ijt}$ is the everyday price consumer i pays for brand j in week t, $Prom_{ijt}$ is a dummy variable for the presence of a price promotion, and $Dscnt_{ijt}$ denotes the value of the price promotion. The consumer's inventory of brand j at the beginning of week t, Inv_{ijt}, is given by

$$(6)\qquad Inv_{ijt} = Inv_{ij(t-1)} + q_{ijt} - C_{ijt}.$$

The variable q_{ijt} is the purchase quantity during period t. We calculate consumption, C_{ijt}, on the basis of previous research (Gupta 1988; see also Ailawadi and Neslin 1998), except we keep track of consumption and inventory at the brand level. We first calculate C_{it}, total category consumption in each week, as $C_{it} = \min(\Sigma_{j=1}^{J} Inv_{ijt}, \bar{C}_i)$, where $\bar{C}_i$ is consumer i's average weekly consumption calculated from the data (Ailawadi and Neslin 1998; Gupta 1988). The calculation shows that the consumer will consume $\bar{C}_i$ as long as he or she has that much inventory on hand. To calculate C_{ijt}, we assume that if the consumer has positive inventory of various brands, he or she consumes the brands in order of preference (ψ_{ijt}). Given that Inv_{ijt} is the inventory of brand j at time t, $\delta_{1i}\Sigma_{j=1}^{J} Inv_{ijt}$ measures total holding cost; δ_{1i} denotes unit holding cost. In addition, $I(\Sigma_{j=1}^{J} Inv_{ijt} - \bar{C}_i)$ is an indicator function equal to 1 when $\Sigma_{j=1}^{J} Inv_{ijt} < \bar{C}_i$ and equal to 0 otherwise. This captures stockout costs (for a more thorough treatment, see Erdem, Imai, and Keane 2003). The combination of Equations 2, 4, and 5 yields the following:

$$(7)\qquad U_{it} = \sum_{j=1}^{J}\left[\psi_{ijt}C_{ijt} - \alpha_{1i}\left(Price_{ijt} - Prom_{ijt}Dscnt_{ijt}\right)d_{ijt}\right] - \alpha_{1i}\delta_{1i}\sum_{j=1}^{J}Inv_{ijt} - \alpha_{1i}\delta_{2i}I\left(\sum_{j=1}^{J}Inv_{ijt} - \bar{C}_i\right) + \alpha_{1i}y_{it}.$$

We add a few adjustments to Equation 7. First, note that α_{1i} represents price response. Previous researchers have found that response to changes in regular price is different from response to promotional price discounts (Guadagni and Little 1983; Mulhern and Leone 1991). We allow for the possibility of having differential response to price promotions and regular price by introducing a new parameter, α_4. Second, we set $-\alpha_{1i}\delta_{1i} = \alpha_{5i}$, where α_{5i} represents the change in utility from increasing inventory by one unit. Third, we set $-\alpha_{1i}\delta_{2i} = \alpha_{6i}$, where α_{6i} represents the change in utility if the consumer is out of stock. Fourth, note that choices are mutually exclusive so that $\Sigma_{j=0}^{J} d_{ijt} = 1$, and when no category purchase is made, $U_{it} = \Sigma_{j=1}^{J}\psi_{ijt}C_{ijt} + \alpha_{5i}\Sigma_{j=1}^{J} Inv_{ijt} + \alpha_{6i}I(\Sigma_{j=1}^{J} Inv_{ijt} - \bar{C}_i) + \alpha_{1i}y_{it}$. Therefore, y_{it}, which does not change across alternatives, is independent of d_{ijt} and will not affect brand choice or the category decision. We therefore treat it as a constant and set it equal to zero (Erdem, Imai, and Keane 2003)[5] and thus can state our utility function as follows:

$$(8)\qquad U_{it} = \sum_{j=1}^{J}(\psi_{ijt}C_{ijt} - \alpha_{1i}Price_{ijt}d_{ijt} + \alpha_{4i}Prom_{ijt}Dscnt_{ijt}d_{ijt}) + \alpha_{5i}\sum_{j=1}^{J}Inv_{ijt} + \alpha_{6i}I\left(\sum_{j=1}^{J}Inv_{ijt} - \bar{C}_i\right).$$

We model promotion availability and consumer perceptions of promotion availability. As do Gönül and Srinivasan (1996) and Assunção and Meyer (1993), we assume that promotion of each brand-size follows a first-order Markov process and is independent across alternatives:[6]

$$(9)\qquad Prob(Prom_{jt} = 1|Prom_{j(t-1)} = 1) = \pi_{j1}, \text{ and}$$

$$(10)\qquad Prob(Prom_{jt} = 1|Prom_{j(t-1)} = 0) = \pi_{j0},$$

for j = 1, ..., J, where π_{j1} denotes the probability of promotion in period t, given that there was promotion in period t − 1 for brand-size j. Similarly, π_{j0} denotes the probability of promotion in period t, given that there was no promotion in period t − 1. The parameters describe the frequencies and the

[4]Note that we include inventory costs in the budget constraint, which is equivalent to including it directly in the utility function. For a similar treatment, see Erdem, Imai, and Keane (2003).

[5]We can relax this assumption and allow total shopping expenditure to affect purchase decision by assuming a nonlinear utility function. We leave modeling expenditure effect for further research.

[6]We checked this assumption by calculating and comparing Equations 9 and 10 and then doing the same using $Prom_{j(t-2)}$ instead of $Prom_{j(t-1)}$. We found that the t − 1 conditionals were different whereas the t − 2 conditionals were virtually the same. For example, for the Heinz 14 oz. bottle, $Prob(Prom_{jt} = 1|Prom_{j(t-1)} = 1) = .02$, whereas $Prob(Prom_{jt} = 1|Prom_{j(t-1)} = 0) = .06$ (the chance of a promotion in week t is greater if the brand-size has not just been promoted). The second-order conditionals for this brand-size are $Prob(Prom_{jt} = 1|Prom_{j(t-2)} = 1) = .05$ and $Prob(Prom_{jt} = 1|Prom_{j(t-2)} = 1) = .05$, or virtually identical. Similar patterns hold for the other brand-sizes.

temporal correlation of promotions over time. We assume that consumers learn these probabilities; Gönül and Srinivasan (1996), who also estimate a structural model, find evidence that consumers can do so, as do Krishna, Currim, and Shoemaker (1991), who use a consumer survey.

Note that the utility function (Equation 8) and perceptions of promotion availability imply that loyal consumers are especially likely to adjust purchase timing (e.g., by acceleration). Consider the consumer who enters the store in time t and finds Brand A on promotion. The consumer does not need to buy the product category now (inventory is high) and realizes that doing so only adds inventory costs in the future. However, the consumer realizes that because the brand is on promotion now, it may not be next week, and by buying now, the consumer can ensure future consumption of Brand A. This future consumption benefit is especially high for consumers who prefer Brand A, and that is why they are intrinsically more likely to adjust purchase timing. Timing adjustment by loyal purchasers is the key intuitive reason for the upward bias in switching estimates; in taking this phenomenon into account, the dynamic structural model should provide more accurate switching estimates.

Our focus is the impact of promotion on brand switching, so it is natural that we model consumer promotion expectations rather than price expectations. Our emphasis on promotion expectations is also based on the following reasons: First, modeling price expectations would add more complexity by increasing the size of the dynamic program state space. Second, including these expectations would only make our case stronger, because there would be even more consumer adjustment of purchase timing. Third, the three price-related quantities are regular price (Price_{ijt}), availability of a promotion (Prom_{ijt}), and depth of a promotion (Dscnt_{ijt}). We believe that consumers are most likely to form expectations for the availability of a promotion. Krishna, Currim, and Shoemaker (1991) find that consumers are more likely to have opinions of promotion frequency than promotion depth. Fourth, the variation in regular price per ounce was fairly small in our data (mean and standard deviations are \$.0561 and \$.0015, \$.0389 and \$.0017, \$.0461 and \$.0015, \$.0480 and \$.0016, \$.0274 and \$.0010, and \$.0349 and \$.0014 for the six brand-size combinations in our application). So plus or minus two standard deviations is on the order of an 8% change, whereas the average price promotion discount is on the order of a 15% to 30% change. Finally, we assume that it is easier to learn the probability of promotion on the next occasion conditional on current promotion than to learn the likelihood of a change in regular shelf price. Therefore, it is more rational to make the effort to learn when promotions will be available because it is easier and the payoff is higher.

The simplest way to incorporate future prices when solving consumers' dynamic programs is to treat the prices constant over the time horizon. Gönül and Srinivasan (1996) follow this approach. We instead represent the prices as random draws from a normal distribution with mean and variance specific to the brand. The mean and variance remain constant over time. The results we report in this article use this specification. We also follow a similar approach for depth of promotions. We observe that constant prices yield similar parameter estimates and identical managerial implications.

The solution to the dynamic program defined by Equation 1 and the specification of our utility function is such that at any time period and given any state, the optimal solution is the solution to the dynamic program from that time forward. The consumer's optimal choice is thus the solution to the Bellman equation:

$$(11)\qquad V_{it}\left(\text{Last}_{it}, \text{Inv}_{it}, \text{Prom}_{it}\right) = \max_{d_{ijt}} E\left\{\sum_{\tau=t}^{T} \delta^{\tau-t} U_{i\tau}\right\}$$

$$= \max_{d_{ijt}}\left\{U_{it} + \delta E_t\left[V_{i(t+1)}\left(\text{Last}_{i(t+1)}, \text{Inv}_{i(t+1)}, \text{Prom}_{i(t+1)}\right)\right]\right\}.$$

There are three state variables: last purchase, inventory, and promotion; Last_{it}, Inv_{it}, and Prom_{it} are vectors that denote consumer i's previous choice, inventory, and current promotion status for all brands, respectively. The value function V_{it} depends on the state at time t for consumer i. It is the maximum expected discounted consumption utility for a consumer who makes decisions about purchase incidence and brand choice given Inv_{it} units of current inventory, last purchase Last_{it}, and observed promotion Prom_{it}.

Heterogeneity and Estimation

Previous research suggests that consumers are heterogeneous in their preferences, sensitivities to marketing-mix variables, holding costs, and stockout costs (e.g., Allenby and Rossi 1991; Chang, Siddarth, and Weinberg 1999; Gönül and Srinivasan 1993; Kamakura and Russell 1989; Krishna, Currim, and Shoemaker 1991). It is also known that ignoring unobserved consumer heterogeneity may lead to biased parameter estimates (Heckman 1981). Let φ_i denote a vector of length J + 5 that represents consumer i's utility parameters $(\alpha_{1i}, \alpha_{2ij}, \alpha_{3i}, \alpha_{4i}, \alpha_{5i}, \alpha_{6i})^T$ for j = 1, …, J, where T denotes the transpose. Then, we assume that φ_i is multivariate normally distributed with the following mean vector and covariance matrix:

$$(12)\qquad \varphi_i \sim N(\varphi_0, \Sigma_\varphi),$$

where $\varphi_0 = (\alpha_1, \alpha_{21}, \ldots, \alpha_{2J}, \alpha_3, \alpha_4, \alpha_5, \alpha_6)$ is the mean of φ_i, and $\Sigma\varphi$ is a variance–covariance matrix of dimension J + 5 in which the diagonal elements denote the corresponding variance of each parameter. Thus, we need to estimate the mean and variance for each of these parameters. In addition, we capture the potential for loyal users to be more promotion sensitive by allowing the intrinsic preference α_{2ij} and promotion coefficients (α_{4i}) to be correlated with correlation ρ. Thus, all the off-diagonal elements of $\Sigma\varphi$ except ρ are constrained to be zero. A positive (negative) ρ indicates that consumers who like the brand are more (less) promotion sensitive than average. Let $\theta = (\alpha_1, \alpha_{2j}, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \sigma_{\alpha 1}, \sigma_{\alpha 2j}, \sigma_{\alpha 3}, \sigma_{\alpha 4}, \sigma_{\alpha 5}, \sigma_{\alpha 6}, \rho, \pi_{j1}, \pi_{j0})$ for j = 1, …, J represent all the parameters to be estimated. We then have the probability of choosing brand j conditional on φ_i and θ ($P_{ijt}[\theta, \varphi_i]$):

$$(13)\qquad P_{ijt}(\theta, \varphi_i) = \text{Prob}(d_{ijt} = 1 | \theta, \varphi_i).$$

We assume that the value function of choice j is distributed by an i.i.d. error term ξ_{ijt}, which follows an extreme value distribution. The probability of consumer i making a sequence of purchase decisions d_{ijt}, j = 0, …, J is given by

(14) $$P_i(\theta, \varphi_i) = \prod_t \prod_j d_{ijt} P_{ijt}(\theta, \varphi_i).$$

Integration over the distribution of φ_i yields the likelihood of observing consumer i's sequence of purchase decisions:

(15) $$P_i(\theta) = E\left[P_{ijt}(\theta, \varphi_i)|\theta\right] = \int_\Phi \prod_t \prod_j d_{ijt} P_{ijt}(\theta, \varphi_i) f(\varphi_i|\theta) d\varphi_i,$$

where Φ is the domain of the previous integration, and $f(\varphi_i|\theta)$ is the multivariate normal probability distribution function for φ_i conditional on θ. Thus, the log-likelihood function to be maximized is

(16) $$\log L(\theta) = \sum_{i=1} \log[P_i(\theta)].$$

There is no closed-form solution for the underlying purchase probabilities in Equation 13. In addition, calculating Equation 15 requires computation of multiple integrals. The dimension of the integrals is too large for traditional numerical integration methods; instead, we use simulated maximum likelihood, which employs Monte Carlo methods to simulate the integrals rather than evaluate them numerically (Keane 1993; McFadden 1989; Pakes 1986). Because inventory is continuous, we have the problem of a large state space. We adopt Keane and Wolpin's (1994) interpolation method by calculating the value functions for a few state space points and using these to estimate the coefficients of an interpolation regression. We then use the interpolation regression function to provide values for the expected maxima at any other state points for which values are needed in the backward recursion solution process.

The adoption of simulation and interpolation significantly reduces the computation burden, which also enables us to take into account unobserved consumer heterogeneity. We begin the simulation with a starting value for θ. Next, we draw a set of parameters for a given consumer. We then solve the dynamic programming problem for the entire time span. We repeat the process for that same consumer, with a different set of φ_i. This generates P_{ijt}. We then go through this process for the next consumer. After all consumers are finished, we have P_{ijt} for each consumer, and we can calculate the likelihood function. Using simulation, we then numerically calculate gradients and obtain a new θ. We repeat the entire process until we can no longer improve the likelihood function.[7]

Reduced-Form Models

We compare our structural model with several conventional reduced-form models that are used to study consumer response to promotion. Using a multinomial logit model, we model the conditional probability that a household chooses brand j. The probability of choosing brand j given that a purchase is made is

(17) $$P_{it}(d_{ijt} = 1|d_{i0t} = 0) = \frac{e^{W^*_{ijt}}}{\sum_{j=1}^J e^{W^*_{ijt}}},$$

where

(18) $$W^*_{ijt} = -\alpha_{1i}\text{Price}_{ijt} + \alpha_{2ij} + \alpha_{3i}\text{Last}_{ijt} + \alpha_{4i}\text{Prom}_{ijt}\text{Dscnt}_{ijt}.$$

Note that Equation 18 corresponds to the structural model in its specification of intrinsic preference, state dependence, price, and price promotion. However, it does not consider purchase incidence or dynamic decision making.

Nested logit is one method for modeling purchase incidence, though it does not explicitly model optimal dynamic decision making. In this framework, the probability of category purchase incidence depends on the expected maximum utility from the brand-choice decision. This expected maximum utility is given by CatVal $= \log(\Sigma_{j=1}^J e^{W^*_{ijt}})$, which reflects the established brand preferences and marketing activity in the category at purchase occasion t (Ben-Akiva and Lerman 1985). Purchase incidence is then represented as

(19) $$P_{it}(d_{i0t} = 0) = \frac{e^{R^*_{it}}}{1 + e^{R^*_{it}}}, \text{ and}$$

(20) $$R^*_{it} = \beta_{0i} + \beta_{1i}C_{it} + \beta_{2i}\text{CatVal}_{it} + \beta_{3i}\sum_{j=1}^J \text{Inv}_{ijt} + \beta_{4i} I\left(\sum_{j=1}^J \text{Inv}_{ijt} - \overline{C}_i\right) + \beta_{5i}\text{PromTime}_{it}.$$

We define inventory and consumption as we did previously. The variable PromTime_{it} represents the time since the last promotion. It is meant to capture that consumers may hold out until the next promotion in a reduced-form sense (Mela, Jedidi, and Bowman 1998).[8] To calculate PromTime_{it}, we calculate the average time between promotions in the category, which is 1.6 weeks. If the time since the latest promotion in the category seen by consumer i is greater than half the average, PromTime_{it} equals 1; otherwise, it equals 0. Given this definition, we expect β_{5i} to be negative. The likelihood function for this nested logit model is

(21) $$L = \Pi_i^I \Pi_j^J \Pi_t^T \left(\frac{e^{W^*_{ijt}}}{\sum_j^J e^{W^*_{ijt}}}\right)^{D_{ijt}} \left(\frac{1}{1 + e^{R^*_{ijt}}}\right)^{1-D_{it}} \left(\frac{e^{R^*_{ijt}}}{1 + e^{R^*_{ijt}}}\right)^{D_{it}},$$

where D_{ijt} is a dummy variable that equals 1 if brand j is chosen by consumer i in week t, and D_{it} is a dummy variable that equals 1 if the product category is purchased by consumer i on trip t.

We examine two additional models of incidence and choice. First is a logit model with no-purchase as an alter-

[7]Note that taking into account unobserved heterogeneity assuming continuous heterogeneity distribution requires the dynamic programming problem to be solved for each draw. A simpler treatment of heterogeneity is latent class approach, as Erdem, Imai, and Keane (2003) adopt.

[8]Note that including the variable PromTime in the nested logit model enables consumers to consider a next-period promotion possibility when making a purchase decision that maximizes their current utility. It is far too simple and ad hoc to capture the complicated decision process generated by forward-looking behavior and its interaction with promotion expectations.

native (Chiang 1991; Erdem and Keane 1996; Erdem and Sun 2002). Our motivation for considering this model is to make the purchase incidence model more similar to our model, which also treats no-purchase as a distinct alternative. Second, we estimate a static version of our model. This assumes the discount factor $\delta = 0$, so consumers do not consider future periods. Accordingly, we also do not include inventory and stockout costs because they require dynamic optimization on the part of the consumer.

ANALYSES

Synthetic Data Simulation

In our first analysis, we simulate synthetic data using the structural model with a known set of parameters, and we estimate various logit models and the structural model on these data. We intend to show that if consumers follow a rational purchase strategy, conventional logit models overestimate brand switching due to promotion.

We generate the data from our proposed dynamic structural model for 200 consumers during a 50-week horizon. In essence, we assume that the structural model is the "true" model. There are two brands with market shares of 70% and 30%. Table 1 shows the true parameter values for generating the synthetic data. We compare five models: (1) a logit choice model with brand preference, price, and promotion as independent variables; (2) a nested logit model; (3) a logit model with no purchase as a choice alternative; (4) the static version of our dynamic model;[9] and (5) the dynamic model.[10]

Table 1, Column 2, shows the true parameter values (mean and standard deviations) for the synthetic data. Columns 3–7 report the average mean and average standard deviations over 100 simulations for each parameter. Table 1 shows the following: First, the structural model is estimated correctly; the average estimated parameters in Column 7 correspond to the true parameters. Second, the nonstructural models consistently overestimated the coefficient for promotion. The upward bias is highest for the logit model and lower for the three choice/incidence models. The results conform to our expectations but do not clearly demonstrate an upward bias in switching elasticities, which can be equal even if the raw coefficients are unequal (Ailawadi, Gedenk, and Neslin 1999).

We calculate promotion-switching elasticity as the percentage change in purchase probability of brand j with and without promotion, conditional on category purchase, for consumer i for brand j at week t:

$$\text{(22)} \qquad e^{s}_{ijt} = \Big[\text{Prob}\big(d_{ijt} = 1 \big| d_{i0t} = 0, \text{Prom}_{ijt} = 1\big) - \text{Prob}\big(d_{ijt} = 1 \big| d_{i0t} = 0, \text{Prom}_{ijt} = 0\big)\Big] \Big/ \text{Prob}\big(d_{ijt} = 1 \big| d_{i0t} = 0, \text{Prom}_{ijt} = 0\big).$$

We calculate this quantity for each household for brand j for an arbitrarily selected Week 8, a week when Brand 1 was not promoted. We simulate promotion availability and use average price discount to represent present promotion depth. We use 100 random draws to simulate consumer heterogeneity. Note that to make the promotion elasticities comparable across models, we calculate the preceding equation for the same consumers for each model, specifically the consumers who make a category purchase in Week 8 in the no-promotion case. In this way, we focus only on the brand-switching aspects of the various models. To complement the switching elasticity, we also calculate the percentage of the immediate sales "bump" that is due to switching rather than acceleration or deceleration.[11] This quantity is comparable to that calculated by other researchers who investigate the switching effect of promotions (e.g., Bell, Chiang, and Padmanabhan 1999; Gupta 1988). Using the same simulation method as we do to calculate the switching elasticity, we classify promotion-week incremental purchases as brand switches (A), accelerated purchases (B), and decelerated purchases (C). The switching percentage is thus A/(A + B + C).

Table 1 presents the switching and switching percentage results and shows the following: First, logit models overestimate promotion elasticities more than the structural model. The t-tests show that the differences are significant. We use the word "overestimate" because the structural model is the true model, we estimate its coefficients accurately with that model, and its elasticities are therefore the true elasticities. Second, all the reduced-form models overestimate the brand-switching effect. The upward bias is highest for the stand-alone logit choice model and second highest for the nested logit model with purchase incidence. Logit with no-purchase options and the static structural models also result in higher switching elasticities.

The results suggest that logit choice models overestimate the brand-switching effect of promotions. The addition of purchase incidence helps alleviate this bias but not completely, apparently because the incidence models are a reduced form of dynamic decision making. Therefore, such a model does not fully capture the consumer's rational purchase strategy and confounds switching with purchase timing.

Empirical Application

Data. We used ketchup data collected by ACNielsen.[12] The calibration sample consists of 8823 observations of 173 households that made 1473 purchases of ketchup during 51 weeks from 1986 to 1988 in Sioux Falls, S.Dak. We excluded occasional purchasers by randomly selecting households that made more than 4 purchases of ketchup during the 51 weeks. We also only chose households that purchased at most one unit of ketchup on each purchase occasion; single-unit purchasers represent 98% of all households

[9]This is a structural model that assumes immediate utility maximization. Consumers do not take into account future utilities. We therefore assume the discount parameter $\delta = 0$ and that there are no inventory effects.

[10]Note that to simplify, we assumed for the simulation that consumers know future regular prices. This assumption is relaxed in our empirical application, in which we assume that consumers view future regular prices as random draws from a normal distribution with constant mean and variance over time.

[11]Note that we cannot calculate this quantity for the stand-alone logit models or the static structural model because these models either do not consider purchase displacement (logit) or do not enable consumers to accelerate or decelerate purchases (static structural). However, it provides a useful basis for comparing the nested logit and no-purchase option logit with the dynamic structural model.

[12]A reason for using ketchup data is that the consumption rate is relative insensitive to promotion, as Ailawadi and Neslin (1998) show. However, for other product categories, ignoring flexible consumption rate may also cause overestimation of brand-switching elasticities. This supports our findings.

Table 1
SIMULATION BASED ON SYNTHETIC DATA

Parameters	*True Values*[a]	*Logit*[a] *(1)*	*Nested Logit (2)*	*Logit with No-Purchase Alternative (3)*	*Static Structural Model (4)*	*Dynamic Structural Model (5)*
Price $-\alpha_1$	−2.00	−3.40 (.34)	−2.68 (.16)	−2.61 (.22)	−2.30 (.19)	−2.22 (.21)
	1.00	1.26 (.27)	1.34 (.44)	1.10 (.44)	1.03 (.25)	1.01 (.26)
Brand preference α_2						
Brand 1	−.50	2.51 (.29)	1.57 (.41)	−.55 (.14)	−.55 (.20)	−.54 (.19)
	.50	1.16 (.24)	.91 (.49)	.56 (.15)	.54 (.09)	.52 (.07)
Brand 2	−1.00			−1.08 (.26)	−1.07 (.20)	−1.07 (.18)
	.50			.51 (.19)	.48 (.22)	.50 (.20)
Promotion α_4	1.50	2.77 (.20)	2.05 (.33)	2.09 (.24)	1.74 (.24)	1.51 (.18)
	1.00	1.19 (.21)	1.43 (.25)	1.31 (.35)	.98 (.30)	.94 (.24)
Unit inventory cost α_5	−.10			−.06 (.03)		−.08 (.04)
	.05			.05 (.02)		.05 (.02)
Stockout cost α_6	−.20			−.22 (.10)		−.18 (.07)
	.05			.06 (.02)		.06 (.02)
π_{10}	.50					.55 (.13)
π_{11}	.25					.26 (.11)
π_{20}	.30					.33 (.09)
π_{21}	.15					.16 (.07)
Purchase Incidence						
Category preference β_0			1.24 (.20)			
			.13 (.06)			
Consumption β_1			.69 (.25)	.62 (.18)		
			.25 (.10)	.30 (.09)		
Category value β_2			.58 (.23)			
			.31 (.15)			
Unit inventory cost β_3			−.08 (.03)			
			.03 (.04)			
Stockout cost β_4			−.03 (.02)			
			.02 (.03)			
Marginal Promotion of Brand 1 in Week 8						
Short-term promotional switching elasticities		.351 (.044)	.275 (.027)	.266 (.021)	.237 (.019)	.198 (.015)
Long-term promotional switching elasticities		.164 (.023)	.156 (.027)	.130 (.016)	.121 (.013)	.101 (.011)
Breakdown of Short-Term Promotion Effect						
Brand switching		N.A.	72%	69%	N.A.	56%
Purchase displacement			28%	31%		44%
Increased Promotion Perception (π) of Brand 1 by 50%						
Percentage change of Brand 1 aggregate sales		.222 (.029)	.186 (.025)	.187 (.025)	.099 (.017)	.054 (.016)

[a]Numbers in the first and second lines are the means and standard deviations of each generated parameter.

Notes: N.A. = not applicable. Number of households = 200; number of weeks = 50; number of observations = 10,000; number of simulations = 100. Numbers not in parentheses are estimated parameters; numbers in parentheses are the standard error of the corresponding parameters. The results are obtained on the basis of 100 simulations. Average means, standard deviations, and their standard errors are reported over 100 simulations. The standard deviation of promotion-switching elasticities are reported in parentheses. To test whether the elasticities differences among Models 1, 2, 3, 4, and 5 are statistically significant, we performed t-tests for the first brand. The t-values are 4.12, 4.26, 3.97, 3.86, and 4.06 for short-term marginal promotion elasticities; 4.18, 3.96, 4.21, 4.19, and 4.40 for long-term marginal promotion elasticities; and 4.85, 4.75, 5.36, 4.86, and 5.02 for policy-change elasticities.

and 96% of all purchase occasions. We selected six leading brand-sizes that constitute more than 85% of the (unit) market share: Heinz 14 oz., 28 oz. and 32 oz., 40 oz. and 44 oz., and 64 oz.; Hunt's 32 oz.; and Del Monte 28 oz. and 32 oz. We grouped similar sizes (e.g., 28 oz. and 32 oz., 40 oz. and 44 oz.) together to reduce the computation burden. The six sizes we considered cover more than 98% of the sales of the three brands in our study. Table 2 reports descriptive statistics. The average price per oz. is $.0561, $.0389, $.0461, $.0480, $.0274, and $.0349 for the six brand-sizes. Average category purchase is 5.344 ozs. per week. We reserved an additional 4131 observations of 81 households that made 647 purchase of ketchup during 51 weeks in Springfield, Mo., for holdout sample validation. In estimating the models, we quantified price, price discount, and inventory relative to a 32 oz. bottle. Thus, if a consumer purchased a 44 oz. bottle, there is 1.375 addition to inventory (q_{ijt} = 1.375). This is a matter of rescaling that does not affect the results.

We chose the ketchup category because it is a product that consumers can use flexibly and that has relatively low holding costs. It is also a well-promoted category. As a result, ketchup is a category for which consumers can plan their

Table 2
DESCRIPTIVE STATISTICS OF KETCHUP DATA

Brand Name	*Market Share*	*Mean Price per Ounce*[a]	*Store Special Price*	
			Frequency	*Mean*
Heinz 14 oz. (1)	4.31%	5.61	5.8%	.896
Heinz 28 oz. and 32 oz. (2)	39.68%	3.89	19.0%	1.114
Heinz 40 oz. and 44 oz. (3)	5.14%	4.61	8.5%	1.022
Heinz 64 oz. (4)	3.37%	4.80	10.9%	1.201
Hunt's 32 oz. (5)	13.52%	2.74	12.5%	.481
Del Monte 28 oz. and 32 oz. (6)	19.47%	3.49	18.1%	1.065
	(85.49%)			

[a]In the estimation, we scaled the prices relative to 32 ozs.
Notes: We selected frequent ketchup users who made at least five purchases during the 51 weeks. The average sample category purchase is .167 bottle per week. The average category purchase is .043 bottle per week in the original data from ERIM.

purchasing. Most of the purchases are single unit, and consumption is relatively stable, as Ailawadi and Neslin (1998) demonstrate. Ketchup therefore seems an ideal category, though certainly not an atypical one, for illustrating rational consumer planning.

We operationalized price as regular, everyday shelf price. We follow Abraham and Lodish (1987) to construct the price discount, Dscnt_{ijt}. We defined promo equal to 1 if there was a price discount in that week and equal to 0 otherwise. We defined Last equal to 1 for the brand-size that was chosen most recently and equal to 0 otherwise (see Ailawadi, Gedenk, and Neslin 1999; Seetharaman, Ainslie, and Chintagunta 1999).[13] The δ is fixed at .995 (Erdem and Keane 1996).[14]

Comparative model fit. The fit statistics for both in-sample (calibration) and out-of-sample (holdout) are provided in Table 3. Because choice models (Model 1 and 2) and choice/incidence models (Models 3–7) fit data to different model structures (incidence and/or brand), the overall log-likelihood values and Bayesian information criteria (BICs) are not directly comparable. Thus, we also report the log-likelihood values of the brand choice component for the nested logits (Models 3 and 4) to compare the performance of various reduced-form choice models. In terms of this choice component, model fit improves from Model 1 to Model 4; Model 4 is the best-fitting choice model. We then compared the overall log-likelihood value and BICs for Models 3–7. The overall log-likelihood (absolute) value and BICs are lower for Model 7 than for Models 3–6, indicating that the dynamic model fits data the best.[15] Holdout sample fit statistics support this finding. To better assess the fit of each model, we report predicted average probabilities and counts of each choice alternative in comparison with the sample frequency and sample count of the calibrated sample. In addition, we report the percentage of correctly predicted choice alternatives and Efron's R^2 for both the calibration sample and the holdout sample.[16] These indexes paint the same picture. The comparison of Models 6 and 7 suggests that the static version of our model performs worse than the dynamic version, which suggests that dynamics indeed matter.

Overall, the dynamic structural model fits better than the other models in terms of log-likelihood, BIC, and hit rates for the calibration and holdout samples.[17] This suggests that our proposed model provides a better description for consumer rational purchase behavior in the ketchup category. Models 3–5 underperform the forward-looking model because they try to capture the dynamics of consumer rational strategy by an exogenously imposed statistical relationship. The assumed relationship is too simple and ad hoc to incorporate the complex effects of purchase timing and forward-looking inventory effects. The choice models (Models 1 and 2) ignore purchase timing and cannot completely capture consumer rational purchase strategy. The static structural model (Model 6) performs worse than the dynamic model because the former leaves out all the dynamics.

Estimated coefficients. Table 4 reports the estimation results of the seven competing models; mean parameter estimates are reported in the first line, and the standard deviation estimates across households are reported in the second line. We first focus on the structural model (Model 7). All the mean coefficients are significant and have expected signs. The standard deviations are also significant for all coefficients except that for stockout cost. This provides evidence that consumers are heterogeneous. Brand-consumption preferences for the six brand-size combinations are also significantly estimated, and there are indications of significant heterogeneity in preference. The significance of the last purchase parameter reinforces the

[13]We estimated all the reduced-form models with smoothed use experience ("BLOY," or brand loyalty, as in Guadagni and Little's [1983] study). There is some improvement of log-likelihood value and BIC, but it does not affect the estimation of other parameters significantly. Comparison of the model estimation can be obtained from the authors.

[14]Although the 51-week time horizon seems long, we note that the discounted factor implies that the distant time periods are less important than immediate time periods.

[15]Akaike information criterion is calculated as –lnL + number of parameters, and BIC is calculated as –lnL + (number of parameters)/2 × ln(number of observations).

[16]Efron's R^2 is calculated as follows:

$$1 - \frac{\sum_{i=1}^{I}\sum_{t=1}^{T}\sum_{j=1}^{J}(Y_{ijt} - \hat{F}_{ijt})^2}{\sum_{i=1}^{I}\sum_{t=1}^{T}\sum_{j=1}^{J}(Y_{ijt} - \bar{Y}_{ijt})^2}.$$

It conveys the proportion of the variance of the dependent variable explained by the independent variable and is a more reliable model selection criterion for discrete-choice models (Amemiya 1985).

[17]Note that the holdout sample consists of households in a different city than the calibration sample households, so this predictive validity test is more stringent than predicting the choices of the same households in a different period or of different households in the same period.

Table 3
MODEL FIT COMPARISONS

Parameter			Logit (1)		Logit with Correlation (2)		Nested Logit (3)		Nested Logit with PromTime (4)		Logit with No-Purchase Alternative (5)		Static Structural Model (6)		Dynamic Structural Model (7)	
Fit Statistics																
In-sample[a]																
–LL			2757.7		2729.1		8947.3		8910.4		8937.7		8920.5		8829.4	
							(2569.3)		(2555.2)							
BIC			2815.9		2791.1		9069.9		9042.1		9060.3		9006.8		8988.4	
							(2638.4)		(2617.2)							
Out-of-sample[b]																
–LL			1485.3		1473.7		4881.8		4859.2		4870.4		4868.0		4840.1	
							(1458.4)		(1438.7)							
	Sample Frequency	*Sample Count*	*Average Probability*[c]	*Prediction*	*Average Probability*	*Prediction*	*Average Probability*	*Prediction*	*Average Probability*	*Prediction*	*Average Probability*	*Prediction*	*Average Probability*	*Prediction*	*Average Probability*	*Prediction*
Predicted Frequency																
Heinz 14 oz.	.0084	74	.0092	82	.0090	81	.0080	68	.0082	69	.0082	71	.0081	68	.0082	72
Heinz 28 oz. and 32 oz.	.0954	842	.1051	870	.1041	868	.1030	861	.1027	858	.1010	861	.1017	858	.1000	849
Heinz 40 oz. and 44 oz.	.0100	88	.0112	99	.0110	98	.0107	96	.0107	95	.0108	95	.0106	97	.0102	92
Heinz 64 oz.	.0074	65	.0057	54	.0059	56	.0065	57	.0068	60	.0070	63	.0067	58	.0071	63
Hunt's 32 oz.	.0092	81	.0088	78	.0090	79	.0094	84	.0093	84	.0089	82	.0087	84	.0090	83
Del Monte 28 oz. and 32 oz.	.0332	293	.0381	310	.0380	307	.0339	305	.0338	302	.0335	300	.0346	306	.0339	299
No purchase	.8364	7380					.8285	7352	.8285	7355	.8306	7351	.8296	7352	.8316	7365
			Calibration	*Holdout*	*Calibration*	*Holdout*	*Calibration*	*Holdout*	*Calibration*	*Holdout*	*Calibration*	*Holdout*	*Calibration*	*Holdout*	*Calibration*	*Holdout*
Hit Rate																
Correctly predicted			.891	.873	.904	.880	.952	.935	.954	.941	.953	.940	.954	.941	.973	.970
Efron's R^2			.041	.039	.046	.043	.052	.050	.054	.051	.053	.050	.055	.053	.064	.064

[a]Number of households = 173; number of weeks = 51; number of observations = 8823. There are 1473 purchases.

[b]Number of households = 81; number of weeks = 51; number of observations = 4131.

[c]To make the average probabilities comparable across all models, for logit models we report the product of average predicted probabilities and percentage of purchases out of the entire sample. For example, .0092 = .05627 × .1635. The average probability for Heinz 14 oz. is .05627, predicted by the logit model when only 16.75% of the 8823 observations are purchase occasions and are considered.

Notes: Because logit models and nested logit models fit different data sets, the overall log-likelihood values are not comparable across these models. Therefore, we report the log-likelihood values of brand choices conditional on purchase occasion and those of purchase occasions for the two nested logit models.

existence of state dependence in brand preference. The results also show that price has a negative impact on purchase, and promotion increases the attractiveness of the price discount, as we expected.

The coefficients for inventory and stockout costs are negative and significant, as we expected. In monetary terms, the holding cost is equivalent to a weekly price reduction of .05/.44 = \$.11. Given an average price of approximately \$1.25 per bottle, a consumer would be indifferent between a 20% price cut (\$.25) and holding a 32 oz. bottle of ketchup in inventory for slightly longer than two weeks. This shows much flexibility in that the consumer would be willing to stockpile a bottle of ketchup even if it meant it was not going to be consumed immediately.

The estimates of the promotion expectations are sensible. The estimated perceived conditional probability of a promotion in period t, given that there was not a promotion in period t – 1, is greater than the probability if there was a promotion in period t – 1. The implied promotion frequencies, calculated as $\pi_{j0}/(1 + \pi_{j0} - \pi_{j1})$, are 7.0%, 19.0%, 9.5%, 7.0%, 12.9%, and 15.7% for the six brand-sizes, respectively, which correspond nicely to the observed frequencies in Table 2 (Heinz 64 oz. is the one exception). Overall, these results are notable: They show that the estimated perceived promotion transition probabilities recover the actual promotion frequencies rather well. This suggests that consumers indeed learn the promotion schedule. The correlation between brand preference and promotion response is positive, showing that consumers who have a higher-than-average intrinsic preference for a particular ketchup brand are also more promotion sensitive than average. Thus, brand-loyal customers are more likely to adjust their purchase timing to meet the promotion schedule of their favorite brand.

We now focus on the purchase incidence parameters in Model 4. The inventory and stockout variables are negative, as we expected (though only weakly statistically significant). The positive estimate of CatVal_{it} shows that consumers are more likely to accelerate purchase on promotion; the negative estimate of PromTime_{it} implies that consumers delay their purchase until next promotion with the belief that the longer the time since the last promotion, the higher is the probability of promotion in the next period. The CatVal_{it} variable is standard for nested logit models, but the PromTime_{it} variable is relatively new. Our results support its inclusion as an ad hoc way to model deceleration.

Finally, the promotion coefficient was consistently higher as we moved from the structural model to Model 5 and through to Model 1. These results are similar to our synthetic data simulation. We next turn to estimated elasticities and switching percentages to determine whether the logit, nested logit, and nondynamic structural models produce greater switching effects.

Estimated elasticities and switching percentages. Table 5 shows the comparative short-term and long-term switching elasticities and percentage breakdown of switching versus purchase displacement for a marginal promotion of Heinz 28 oz. and 32 oz. in Week 11. It also reports sales elasticities when the promotion perceptions of Heinz 28 oz. and 32 oz. are increased by 50%. The results are quite similar to our simulation in that the switching effects are consistently higher for the logit, nested logit, logit with no-purchase option, and static structural models than for the dynamic structural model. The nondynamic models infer more brand switching than does the dynamic structural model. Given our concerns about not taking into account consumer rational timing adjustments and given our previous simulation results, the empirical results suggest that reduced-form models overstate switching effects compared with a dynamic structural model.

It is notable that there is not much difference between the two logit choice models or among the four reduced-form choice/incidence models, though PromTime seems to help. The key seems to be the addition of purchase incidence, but how incidence is added does not make much of a difference, except in the dynamic structural model.

The difference between the reduced-form choice/incidence models and the structural model is managerially meaningful. The difference in short-term brand-switching elasticities between Model 5 and the structural Model 7 is .346 – .242 = .104. That means that Model 5 finds an additional 10.4% change in base probabilities compared with the structural model. Assuming a base of roughly .40 purchase probability for Heinz 28 oz./32 oz. (its market share) means a difference in .104 × .40 = 4.16 share points. A share point in the ketchup category is worth roughly \$3 million dollars in revenue, so the difference between these models is a matter of \$12.5 million, or \$240,000 per week.[18] The difference between the logit choice models and the structural model is .426 – .242 = .184 × .40 = 7.4 share points, or \$22.1 million, or \$424,615 per week. These are managerially meaningful differences (see, e.g., Abraham and Lodish 1987).

Policy-Change Analysis

Senior managers are often more interested in evaluating the effect of a significant change in promotion policy rather than the effect of marginal promotion. It is in this area that we expect the dynamic structural model to distinguish itself even more from the nondynamic models, because the structural model explicitly takes into account the promotion policy through the frequency perception coefficients, π. Keane (1997a, p. 312) states our expectations best: "To successfully forecast behavior after a regime change one needs to model how agents tailor their decision rules to particular regimes. One needs a so-called 'structural' model whose parameters are 'primitives' of agents' preferences, information processing systems,... which remain fixed across regimes."

We expect that the structural model will predict less of a gain in market share than the other models for the brand that experiences a quantum increase in promotion frequency. This is because these models do not explicitly enable consumers to adjust their purchase timing as a function of the firm's overall promotion policy. The choice/incidence models try to achieve this by including inventory and ad hoc variables such as PromTime_{ijt}, but they do not do as complete a job as the structural model, which models rational consumer adjustments explicitly in a dynamic programming framework.

We run a simulation in which we increase the actual promotion frequencies of Heinz 28 oz./32 oz. by 50% and

[18]Source: Information Resources (1995). Volume of mustard and ketchup per 1000 households is 5916 units; ketchup accounts for 77% of the category and is priced at \$.73 per unit. The assumption of 90 million households translates to a category sales volume of \$5,916 × .77 × .73 × 90,000 = \$299.3 million.

Table 4
MODEL ESTIMATION WITH KETCHUP DATA

Parameter	*Logit (1)*	*Logit with Correlation (2)*	*Nested Logit (3)*	*Nested Logit with PromTime (4)*	*Logit with No-Purchase Alternative (5)*	*Static Structural Model (6)*	*Dynamic Structural Model (7)*
Price $-\alpha_1$	–.61 (.18)	–.59 (.12)	–.51 (.14)	–.46 (.12)	–.47 (.14)	–.40 (.08)	–.44 (.08)
	.23 (.10)	.20 (.09)	.19 (.06)	.17 (.06)	.16 (.07)	.10 (.06)	.11 (.04)
Brand preference α_2							
Heinz 14 oz. (1)	–.12 (.03)	–.14 (.05)	–.17 (.07)	–.16 (.06)	–1.05 (.22)	–.97 (.13)	–.94 (.16)
	.01 (.002)	.02 (.002)	.05 (.01)	.05 (.002)	.33 (.10)	.29 (.08)	.34 (.11)
Heinz 28 oz. and 32 oz. (2)	1.48 (.27)	1.43 (.21)	1.51 (.32)	1.49 (.29)	–.14 (.08)	–.20 (.06)	–.17 (.08)
	.62 (.14)	.60 (.12)	.57 (.12)	.54 (.13)	.06 (.02)	.04 (.01)	.05 (.01)
Heinz 40 oz. and 44 oz. (3)	–.05 (.009)	–.05 (.01)	–.08 (.02)	–.07 (.02)	–.86 (.29)	–.80 (.16)	–.70 (.21)
	.02 (.005)	.02 (.006)	.04 (.01)	.04 (.01)	.23 (.08)	.20 (.06)	.20 (.06)
Heinz 64 oz. (4)	–.15 (.04)	–.17 (.05)	–.19 (.06)	–.21 (.05)	–1.10 (.33)	–1.17 (.22)	–1.13 (.20)
	.06 (.02)	.07 (.03)	.09 (.04)	.01 (.04)	.28 (.08)	.22 (.07)	.20 (.09)
Hunt's 32 oz. (5)	–.39 (.12)	–.37 (.11)	.61 (.08)	.63 (.07)	–.50 (.10)	–.51 (.13)	–.49 (.12)
	.20 (.11)	.19 (.11)	.25 (.07)	.27 (.09)	.17 (.07)	.15 (.10)	.15 (.10)
Del Monte 28 oz. and 32 oz. (6)					–.64 (.29)	–.58 (.22)	–.50 (.24)
					.40 (.14)	.30 (.10)	.34 (.10)
Last purchase α_3	2.40 (.31)	2.36 (.32)	1.89 (.22)	1.86 (.19)	1.76 (.15)	1.41 (.26)	1.31 (.22)
	.67 (.19)	.54 (.18)	.45 (.15)	.58 (.22)	.55 (.09)	.98 (.29)	.98 (.30)
Promotion α_4	.33 (.07)	.31 (.07)	.24 (.06)	.22 (.08)	.19 (.07)	.17 (.06)	.09 (.05)
	.16 (.07)	.11 (.03)	.09 (.04)	.08 (.05)	.08 (.03)	.08 (.03)	.10 (.02)
Unit inventory cost α_5					–.06 (.02)		–.05 (.02)
					.010 (.007)		.018 (.005)
Stockout cost α_6					–.11 (.07)		–.25 (.08)
					.09 (.08)		.16 (.10)
Correlation between brand preference and promotion ρ		.15 (.09)	.13 (.10)	.13 (.08)	.17 (.07)	.16 (.07)	.18 (.09)
π_{10}							.074 (.01)
π_{11}							.011 (.03)
π_{20}							.209 (.09)
π_{21}							.110 (.05)
π_{30}							.099 (.03)
π_{31}							.054 (.04)
π_{40}							.071 (.03)
π_{41}							.056 (.03)
π_{50}							.141 (.04)
π_{51}							.049 (.03)
π_{60}							.173 (.06)
π_{61}							.072 (.07)
Purchase Incidence							
Category preference β_0			.25 (.07)	.27 (.11)			
			.12 (.05)	.11 (.04)			
Consumption rate β_1			.52 (.21)	.54 (.19)	.36 (.07)		
			.23 (.07)	.22 (.08)	.17 (.05)		
Category value β_2			.18 (.06)	.19 (.07)			
			.07 (.04)	.09 (.05)			
Inventory β_3			–.04 (.03)	–.05 (.03)			
			.03 (.02)	.04 (.02)			
Stockout β_4			–.08 (.09)	–.06 (.07)			
			.04 (.13)	.03 (.08)			
PromTime β_5				–.06 (.02)	–.08 (.02)		
				.04 (.02)	.03 (.01)		

Notes: For each parameter, the top number not in parentheses is the estimated mean value of the parameter; the number beneath it that is not in parentheses is the estimated standard deviation of the parameter. Numbers in parentheses are the standard errors of the various estimates.

assume that perceived promotion frequencies also increase by that amount in the dynamic structural model (π_{20} and π_{21}).[19] Given that the base promotion frequency is 19% for Heinz 28 oz./32 oz., we randomly select 10% more of the weeks when there was no promotion, assume that there were promotions running in those weeks, and calculate the predicted percentage change in Heinz 28 oz./32 oz. sales. The results in Table 5 confirm our expectations. The logit and nested logit models all predict substantially greater increases in sales than the dynamic structural model, and the results are even more dramatic than those for a marginal additional promotion. The predicted sales changes for Heinz 28 oz./32 oz. are more than double for the nonstructural models compared with the dynamic structural model. For example, the nested logit model with PromTime predicts a 19.1% increase in sales, whereas the structural model predicts only an 8.5% increase.

The results suggest that reduced-form models significantly overestimate the effectiveness of a policy increase in promotion frequency because they do not explicitly enable consumers to adjust their purchase timing strategies rationally in response to the new regime, and they assume the same promotion sensitivities hold for the new regime.

Limitations

As we noted in the model development section, we assume a simple mechanism for consumers' future price expectations. We focused on expectations of promotions; notably, regular price variations were quite modest (see "Model Specification"). Therefore, we resorted to a simplifying assumption to enhance the computational feasibility of our model. Erdem, Imai, and Keane (2003) and Hendel and Nevo (2002) systematically examine the price expectation process. A comprehensive specification of both price and promotion expectations is desirable, particularly when analyzing a category in which price fluctuations are considerable. We hope to address this issue in a future research effort.

We also implicitly assume that consumers expect promotions to be independent across alternatives. Erdem, Imai, and Keane (2003) and Hendel and Nevo (2002) find that the promotions across alternatives may be correlated. A simple pairwise empirical analysis shows that only 5 of the 15 average within-store correlations among our six brand-sizes are significantly different from 0. The magnitude of the significant ones ranges from –.16 to .27. As part of increasing the sophistication of consumer expectations, we might select a higher-order Markov model that allows for covariation in promotions across alternatives. This would complicate the model development significantly; however, it would be a worthy topic for further research. In the interim, we conjecture that any better specification of consumer expectations should strengthen the findings reported here. As consumers hold more sophisticated expectations of the firms' promotional strategies, they will become more adept at adjusting their purchase timing, which increases the need for dynamic rational models to unravel purchase-timing decisions from brand switching.

Finally, we have assumed that the consumption rate is deterministic. As a result, the advent of a stockout is deterministic. This happens because our model predicts the same choices for households with identical purchase histories and holds taste parameters constant. This may result in a bias in elasticities (for details, see Erdem, Imai, and Keane 2003). With six alternatives, the possibility of identical purchase histories should decrease as the purchase history increases. Given the long history for each household in our model, we believe that the bias arising from deterministic consumption should be minimal.

SUMMARY AND DISCUSSION

The themes of this article are that logit decision models can overestimate the extent of promotion-induced brand switching because they do not completely account for consumers' rational adjustments in purchase timing and that a dynamic structural model can address this problem. We have investigated this thesis by developing a dynamic structural model of choice and incidence and comparing it with a stand-alone logit choice model, a nested logit model with incidence, a logit model with a no-purchase alternative, and a static version of the dynamic model. The dynamic model differs from these reduced-form static models by explicitly modeling the process by which consumers decide what and when to purchase, taking into account their current inventory status and their perceptions of future promotion activity. Our key findings are as follows:

- In a simulation that uses data generated by the structural model, the reduced-form models consistently overpredicted switching elasticities and the percentage of the promotion sales bump due to brand switching.
- The structural model fits the data better than the reduced-form models, including the reduced-form choice/incidence models.
- In estimating these models on real data for the ketchup category, we observed the same pattern of estimated elasticities and switching percentages that we found in the simulation.
- In evaluating the total effect on sales of a 50% increase in promotion frequency, the reduced-form models predicted much higher sales gains than the dynamic structural model.

The results suggest that, as we expected, reduced-form models overstate the brand-switching effects of promotion. All four analyses contribute to this conclusion. For example, the simulation convinces that the reduced-form models overstate elasticities on simulated data, but there is no guarantee that the simulated data, based on the structural model, reflect reality. The better fit of the dynamic structural model on real data, together with its sensible coefficient estimates and our observation of the same pattern in estimated elasticities in real data, suggests that the simulation reflects reality well and that the higher switching effects by the logit models in the empirical test are truly overstated. The policy analysis drives home the point that the real gain from the structural model is in evaluating major policy changes, because that is an area in which the model's explicit accounting for the current promotion policy works to advantage.

As we expected, in all our analyses, the stand-alone logit model exhibits the most bias. Much, though not all, of the bias can be alleviated by estimating nested logit models with incidence; at least these models try to attain similar

[19]Note that we assume that consumers learn the new schedule and update their frequency perceptions (π). The consumer may not be perfectly informed of the policy change. Because our article focuses more on how consumers adjust purchase given their perceptions on promotion schedule than on how they form promotion perceptions, we make this assumption and leave perception formation for further research.

Table 5
PROMOTION ELASTICITIES

Scenarios	*Logit (1)*	*Logit with Correlation (2)*	*Nested Logit (3)*	*Nested Logit with PromTime (4)*	*Logit with No-Purchase Alternative (5)*	*Static Structural Model (6)*	*Dynamic Structural Model (7)*
Marginal Promotion of Heinz in Week 11							
Short-term switching elasticity							
Heinz 28 oz. and 32 oz.	.431 (.062)	.426 (.064)	.378 (.070)	.347 (.067)	.346 (.066)	.333 (.052)	.242 (.045)
Long-term switching elasticity							
Heinz 28 oz. and 32 oz.	.137 (.033)	.135 (.031)	.117 (.028)	.108 (.023)	.109 (.026)	.099 (.006)	.074 (.007)
Breakdown of short-term promotion effect							
Brand switching	N.A.	N.A.	79%	74%	77%	N.A.	56%
Purchase displacement			21%	26%	23%		44%
Increased Promotion Perception (π) of Heinz by 50%							
Percentage change of aggregate sales							
Heinz 28 oz. and 32 oz.	.253 (.041)	.235 (.040)	.196 (.026)	.191 (.031)	.190 (.038)	.190 (.040)	.085 (.0017)

Notes: To test whether the elasticities difference between Model m = 1, 2, ..., 6 and Model 7 are statistically significant, we performed t-tests for Heinz. The t-values are 3.49, 3.81, 3.03, 2.69, 2.92, 3.44, and 3.59 for short-term marginal promotion elasticities; 2.88, 2.37, 2.81, 2.08, 2.48, 3.53, and 3.47 for long-term marginal promotion elasticities; and 4.01, 4.06, 3.08, 3.23, 3.57, 3.75, and 3.80 for policy-change elasticities. Standard errors are in parentheses.

phenomena as the structural model by including variables such as PromTime as a means to model consumers' holding out for the next promotion. They can also account for loyal consumers' taking most advantage of promotions. These steps improve the elasticities and sales predictions but not as completely as does the dynamic structural model.

We have also learned the following from the particular coefficients and results of our empirical study:

- Consumers with intrinsically higher preference are more likely to take advantage of a promotion for their preferred brand.
- Consumers appear to have accurate perceptions of promotion frequency in that our estimated parameters for promotion expectations corresponded well to actual promotion frequency.
- Deceleration is a real phenomenon, as exhibited most directly by the significance of the PromTime variable.
- A dynamic structural model of choice/incidence fits the data better and performs better on holdout data testing than do reduced-form models that attempt to cover the same phenomena. This suggests that the effort of running consumers through a dynamic program pays off in better fit and prediction.

Our findings have several important implications for researchers. First, the choice promotion–elasticity estimates derived from reduced-form models, especially stand-alone logit models, should be interpreted with caution. These elasticities are probably overstated. Second, researchers should devote more attention to dynamic structural models. We perceive movement in this direction (Erdem, Imai, and Keane 2003; Erdem and Keane 1996; Gönül and Srinivasan 1996; Sun 2000) and encourage research in this area. The cost of computation for these models is high, but with improved computing power they are becoming more practical. That the benefit of these models is relevant to managers (i.e., more accurate predictions of brand switching and sales levels) should spur the movement even more. Third, in the event that a researcher is unable to estimate a structural model, the next best model is an ad hoc choice/incidence model that includes variables such as PromTime. These variables are statistically important and help improve switching elasticity estimates. Fourth, our results encourage theoretical research that emphasizes purchase dynamics as a reason for promotion (see Blattberg, Eppen, and Lieberman 1981). There is a need for additional economic models that examine the effects of dynamic rational consumer behavior on promotion policy in a competitive framework.

MANAGERIAL IMPLICATIONS, CONCLUSION, AND FURTHER RESEARCH

First, promotions may be more of a purchase-timing game than a brand-switching game, as was previously believed. This is consistent with recent research that has shown the importance of stockpiling and deceleration (Bell, Chiang, and Padmanabhan 1999; Kopalle, Mela, and Marsh 1999; Mela, Jedidi, and Bowman 1998; Seetharaman, Ainslie, and Chintagunta 1998; van Heerde, Gupta, and Wittink 2003; van Heerde, Leeflang, and Wittink 2000).

Second, managers should be cautious of panel data–based simulations of forecasted changes in sales due to major policy changes, if the forecasting models do not explicitly take into account how consumers will adjust to the new policy. Such forecasts may be overoptimistic because they do not consider how consumers will adjust to the policy change. For example, if reliance on promotion is doubled, many consumers who formerly bought at regular price will structure their buying habits to purchase during promotions. These are not truly incremental sales.

Third, because loyal customers are more likely to purchase on promotion by adjusting purchase timing, manufacturers should find ways to distinguish loyal customers from nonloyal customers and target the latter for promotion effort.

Fourth, because consumer rational purchase strategies can have a negative impact on profit, manufacturers should find ways to discourage consumers from taking advantage of promotion by adjusting purchase timing. For example, specifying the maximum number of units purchased on promotion can increase the promotion effect of attracting new purchases switched from other brands or other stores but can limit stockpiling behavior.

Although our results are quite encouraging, there are remaining issues that suggest topics for further research. First, it would be worthwhile to find reduced-form models that capture the incidence and choice process as well as the structural model. We believe this would be possible if the reduced-form models contained the correct variables; for example, promotion frequency could be incorporated directly in the model (see Kopalle, Mela, and Marsh 1999). Second, as we mentioned previously, we can relax the assumption that consumption is constant. For some product categories, promotion also has a significant impact on consumption rate, which contributes to the incremental sales increase that promotion induces (for a structural model in which consumers endogenously decide how much to consume, see Sun 2003). Third, it would be beneficial to examine profit. Promotion causes brand switching, which contributes positively to profit; however, promotion induces purchase acceleration or deceleration and stockpiling, which may contribute negatively to profit. What promotion strategy can balance these considerations in the presence of a consumer who will rationally adjust to any policy change? Fourth, as we also mentioned previously, regular price expectations and promotion depth as well as promotion frequency could be modeled. This would add complexity because of the additional continuous state space variable and is unlikely to change our basic result. However, the added insight of distinguishing between promotion and price would be important. Relatedly, rather than treat the promotion schedules as independent, further research could incorporate dependencies. Fifth, it would be especially useful to model the process by which consumers learn promotion schedules. This would enhance the message from our Lucas critique simulation, which, for the sake of illustration, assumed that consumers readily learned the new promotion schedule. Finally, estimates of brand switching are at least indirectly related to the problem of estimating the "postpromotion dip" (e.g., Neslin and Stone 1996). The structural model we develop, embellished, for example, by incorporating endogenous consumption, the number of units purchased, and transaction costs, would be useful for identifying the key factors that mask the postpromotion dip.

Although we demonstrate the merit of the structural model, we must recognize some of the limitations of the approach. These models pose difficult estimation challenges that necessitate simplifying assumptions. More important, developing and estimating such a model may be beyond the reach of most practitioners. Therefore, managers may continue to rely on reduced-form models that are far simpler to estimate. By developing several reduced-form models for comparison purposes, we offer guidance for specifications that might alleviate, at least partially, the potential bias for overestimating the impact of promotion on brand switching.

In conclusion, this research provides encouraging support that more accurate estimates of brand-switching elasticities can be obtained by incorporating forward-looking consumer behavior into structural models. This calls for further research projects in this area. As the buying public becomes more and more informed about prices and shopping issues, we believe that taking into account the rational consumer (Keane 1997b) will become more important in evaluating the effectiveness of sales promotions.

REFERENCES

Abraham, Magid M. and Leonard M. Lodish (1987), "PROMOTER: An Automated Promotion Evaluation System," *Marketing Science*, 6 (2), 101–123.

Ailawadi, Kusum L., Karen Gedenk, and Scott A. Neslin (1999), "Heterogeneity and Purchase Event Feedback in Choice Models: An Empirical Analysis with Implications for Model Building," *International Journal of Research in Marketing*, 16 (3), 177–98.

——— and Scott A. Neslin (1998), "The Effect of Promotion on Consumption: Buying More and Consuming It Faster," *Journal of Marketing Research*, 35 (August), 390–98.

Allenby, Greg M. and Peter E. Rossi (1991), "Quality Perceptions and Asymmetric Switching Between Brands," *Marketing Science*, 10 (3), 185–204.

Amemiya, Takeshi (1985), *Advanced Econometrics*. Cambridge, MA: Harvard University Press.

Assunção, João L. and Robert Meyer (1993), "The Rational Effect of Price Promotions on Sales and Consumption," *Management Science*, 39 (5), 517–35.

Bell, David R., Jeongwen Chiang, and V. Padmanabhan (1999), "The Decomposition of Promotional Response: An Empirical Generalization," *Marketing Science*, 18 (4), 504–526.

Ben-Akiva, Moshe and Steven R. Lerman (1985), *Discrete Choice Analysis*. Cambridge: Massachusetts Institute of Technology Press.

Blattberg, Robert C., Gary Eppen, and Joshua Lieberman (1981), "A Theoretical and Empirical Evaluation of Price Deals for Consumer Nondurables," *Journal of Marketing*, 45 (January), 116–29.

Bucklin, Randolph E., Sunil Gupta, and S. Siddarth (1998), "Determining Segmentation in Sales Response Across Consumer Purchase Behaviors," *Journal of Marketing Research*, 35 (May), 189–97.

Chang, Kwangpil, S. Siddarth, and Charles B. Weinberg (1999), "The Impact of Heterogeneity in Purchase Timing and Price Responsiveness on Estimates of Sticker Shock Effect," *Marketing Science*, 18 (2), 178–92.

Chiang, Jeongwen (1991), "A Simultaneous Approach to the Whether, What, and How Much to Buy Questions," *Marketing Science*, 10 (4), 297–315.

——— (1995), "Competing Coupon Promotions and Category Sales," *Marketing Science*, 14 (1), 105–122.

Chintagunta, Pradeep K. (1993), "Investigating Purchase Incidence, Brand Choice, and Purchase Quantity Decisions of Households," *Marketing Science*, 12 (2), 184–208.

———, Dipak C. Jain, and Naufel J. Vilcassim (1991), "Investigating Heterogeneity in Brand Preferences in Logit Models for Panel Data," *Journal of Marketing Research*, 28 (November), 417–28.

Erdem, Tülin, Susumu Imai, and Michael P. Keane (2003), "Brand and Quantity Choice Dynamics Under Price Uncertainty," *Quantitative Marketing and Economics*, 1 (1), 5–64.

——— and Michael P. Keane (1996), "Decision-Making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," *Marketing Science*, 15 (1), 1–20.

——— and Baohong Sun (2002), "An Empirical Investigation of the Spillover Effects of Advertising and Sales Promotions in Umbrella Branding," *Journal of Marketing Research*, 39 (November), 408–420.

Golabi, Kamal (1985), "Optimal Inventory Policies When Ordering Prices Are Random," *Operations Research*, 33 (3), 575–88.

Gönül, Fusun and Kannan Srinivasan (1993), "Modeling Multiple Sources of Heterogeneity in Multinomial Logit Models: Methodological and Managerial Issues," *Marketing Science*, 12 (3), 213–29.

——— and ——— (1996), "Impact of Consumer Expectations of Coupons on Purchase Behavior," *Marketing Science*, 15 (3), 262–79.

Guadagni, Peter M. and John D. Little (1983), "A Logit Model of Brand Choice Calibrated on Scanner Data," *Marketing Science*, 2 (3), 203–238.

Gupta, S. (1988), "Impact of Sales Promotions on When, What, and How Much to Buy," *Journal of Marketing Research*, 25 (November), 342–55.

Hanemann, W. Michael (1984), "Discrete/Continuous Models of Consumer Demand," *Econometrica*, 52 (3), 541–61.

Heckman, J. (1981), "Heterogeneity and State Dependence," in *Studies in Labor Markets*, S. Rosen, ed. Chicago: University of Chicago Press, 91–139.

Helsen, Kristian and David C. Schmittlein (1992), "Some Characterizations of Stockpiling Behavior Under Uncertainty," *Marketing Letters*, 3 (1), 5–17.

Hendel, Igal and Aviv Nevo (2002), "Measuring the Implications of Sales and Consumer," working paper, University of California, Berkeley.

Information Resources Inc. (1995), *IRI Marketing Factbook*. Chicago: Information Resources.

Kamakura, Wagner A. and Gary J. Russell (1989), "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure," *Journal of Marketing Research*, 26 (November), 379–90.

Keane, Michael P. (1993), "Simulation Estimation for Panel Data Models with Limited Dependent Variables," in *Handbook of Statistics*, G.S. Maddala, C.R. Rao, and H.D. Vinod, eds. New York: Elsevier Science Publishers.

——— (1997a), "Current Issues in Discrete Choice Modeling," *Marketing Letters*, 8 (3), 307–322.

——— (1997b), "Modeling Heterogeneity and State Dependence in Consumer Choice Behavior," *Journal of Business and Economic Statistics*, 15 (3), 310–27.

——— and Kenneth I. Wolpin (1994), "Solution and Estimation of Dynamic Programming Models by Simulation," *Review of Economics and Statistics*," 76 (4), 648–72.

Kopalle, Praveen K., Carl F. Mela, and Lawrence Marsh (1999), "The Dynamic Effect of Discounting on Sales: Empirical Analysis and Normative Pricing Implications," *Marketing Science*, 18 (3), 317–32.

Krishna, Aradhna (1992), "The Normative Impact of Consumer Price Expectations for Multiple Brands on Consumer Purchase Behavior," *Marketing Science*, 11 (3), 266–87.

——— (1994a), "The Effect of Deal Knowledge on Consumer Purchase Behavior," *Journal of Marketing Research*, 31 (February), 76–91.

——— (1994b), "The Impact of Dealing Patterns on Purchase Behavior," *Marketing Science*, 13 (4), 351–73.

———, Imran C. Currim, and Robert W. Shoemaker (1989), "Consumer Perceptions of Promotional Activity," *Journal of Marketing*, 55 (April), 4–16.

Lucas, Robert, Jr. (1976), "Econometric Policy Evaluation: A Critique," in *The Phillips Curve and Labor Market*, Carnegie-Rochester Conference on Public Policy Series, Vol. 1, Karl Brunner and Allan H. Meltzer, eds. New York: American Elsevier, 19–46.

McFadden, Daniel (1989), "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration," *Econometrica*, 57 (5), 995–1026.

Mela, Carl F., Kamel Jedidi, and Douglas Bowman (1998), "The Long-Term Impact of Promotions on Consumer Stockpiling Behavior," *Journal of Marketing Research*, 34 (May), 248–61.

Meyer, Robert J. and João Assunção (1990), "The Optimality of Consumer Stockpiling Strategies," *Marketing Science*, 9 (1), 18–41.

Mulhern, Francis J. and Robert P. Leone (1991), "Implicit Price Bundling of Retail Products: A Multiproduct Approach to Maximizing Store Profitability," *Journal of Marketing*, 55 (October), 63–76.

Neslin, Scott A., Caroline Henderson, and John Quelch (1985), "Consumer Promotions and the Acceleration of Product Purchases," *Marketing Science*, 4 (3), 147–65.

——— and Linda G. Schneider Stone (1996), "Consumer Inventory Sensitivity and the Postpromotion Dip," *Marketing Letters*, 7 (1), 77–94.

Pakes, Ariel (1986), "Patents as Options: Some Estimates of the Value of Holding European Patent Stocks," *Econometrica*, 54 (4), 755–84.

Seetharaman, P.B., Andrew Ainslie, and Pradeep K. Chintagunta (1999), "Investigating Household State Dependence Effects Across Categories," *Journal of Marketing Research*, 36 (November), 448–500.

Sun, Baohong (2003), "Promotion Effect on Endogenous Consumption," working paper, Kenan-Flagler Business School, University of North Carolina.

van Heerde, Harald J., Sachin Gupta, and Dick R. Wittink (2003), "Is 75% of the Sales Promotion Bump Due to Brand Switching? No, Only 33% Is," *Journal of Marketing Research*, 40 (November), 481–91.

———, Peter S.H. Leeflang, and Dick R. Wittink (2000), "The Estimation of Pre- and Postpromotion Dips with Store-Level Scanner Data," *Journal of Marketing Research*, 37 (August), 383–95.