



To Whom, When, and How Much to Discount? A Constrained Optimization of Customized Temporal Discounts

Joseph Johnson^{a,*}, Gerard J. Tellis^{b,1}, Edward H. Ip^{c,2}

^a University of Miami, United States

^b Marshal School of Business, University of Southern California, United States

^c Wake Forest University School of Medicine, 3rd Floor MRI Building Medical Center Boulevard, Winston-Salem, NC 27103, United States

Abstract

Customized temporal discounts are price cuts or coupons that are tailored by size, timing, and household to maximize profits to a retailer or manufacturer. The authors show how such discounts allow companies to optimize to whom, when, and how much to discount. Such a scheme allows firms to send just enough discounts just prior to the individual's purchase of a rival brand. To do so, the authors model household purchase timing and brand choice in response to discounts and use Bayesian estimation to obtain individual household parameters. They illustrate the model on a Japanese data set having price cuts, a US data set having coupons, and another US data set having discounts. They formulate the optimization task of customized temporal coupons as a constrained multiple-knapsack problem under a given budget. They use simulations of the empirical contexts to obtain optimal solutions and to assess improvement in profits relative to existing practice and alternate models in the literature. The proposed model yields increase in profits of 18–40 percent relative to a standard model that optimizes the value but not timing of discounts. © 2013 New York University. Published by Elsevier Inc. All rights reserved.

Keywords: Customized temporal discounts; Constrained optimization; Marketing budgets

Introduction

Discounts are important tools in the strategic arsenal of firms. Each year firms spend billions of dollars on such price promotions in an attempt to induce consumers to switch to their brands. For example, Catalina Marketing, one distributor of promotions, retrieves about 250 million transactions per week across more than 21,000 grocery stores, analyzes over 100 million households to customize promotions to them and earns \$400 million (<http://www.catalinamarketing.com>).

Salop and Stiglitz (1977) were the first to suggest that price differences across consumers for the same product was discrimination by firms between informed and uninformed consumers. Varian (1980) extended this analysis to the problem of temporal competition among retailers through the use of “sales”. Narasimhan (1984) showed that coupons were a more efficient means of price discrimination than price-cuts because the former

required more effort. By this logic a coupon is just a selectively distributed discount, requiring some kind of action by a buyer. So we use the term discount broadly, to cover any temporary price reduction either as a price-cut good on purchase, or as a coupon or rebate good on redemption.

Nowadays, consumers differ not only in their knowledge of, but also their sensitivity to, prices at various points in time. Thus an optimal price discrimination scheme would be to sell to individual consumers at the time that each is willing to pay the highest price. Such a scheme requires high list price while targeting individual consumers with discounts that just motivate them to switch to the discounted brand. Further, such discounts must arrive at a time just prior to the individual's purchase of a rival brand. We refer to this form of differential pricing timed with personalized discounts as *customized temporal discounts*. It addresses the key question in differential pricing: *To whom to discount, when, and how much*.

Currently, most differential pricing takes the form of differentiating groups of consumers by geographical location, brand loyalty, product bundle, or timing. The revolution in media, data, and methods will enable the adoption of customized temporal discounts – a highly selective form of differentiation at the level of individual customers on particular days, based upon

* Corresponding author. Tel.: +1 305 284 1379.

E-mail addresses: jjohnson@miami.edu (J. Johnson), tellis@usc.edu (G.J. Tellis), eip@wfubmc.edu (E.H. Ip).

¹ Tel.: +1 213 740 5031.

² Tel.: +1 336 716 9833.

consumer response to past promotions. Customized temporal discounts provide marketers with a unique opportunity to exploit what [Varian \(2000\)](#) calls perfectly efficient pricing. The list price minus the customized discount is set to each individual's marginal willingness to pay on a particular day. The use of electronic household data allows an analyst to determine the optimal value and time of the discount for each individual. Online retailers like Peapod or promotion intermediaries like Catalina can then use such temporally customized discounts for perfectly efficient pricing.

Customizing the timing of the discount can be crucial. To see this, consider a price-cut. A price-cut offered too soon may not generate a response from consumers with adequate inventory. A price-cut offered too late may lose sales to competing brands. Offering an early coupon with a late expiration date does not resolve the timing problem because consumers might forget about or misplace a coupon they receive too early or respond to a more recent one from a competitor. These reasons may explain why response to coupons decays exponentially ([Inman and McAlister 1994](#)). Thus timing a discount is important and offering it too soon or too late is suboptimal.

Despite the importance of optimizing the value and timing of discounts, no study in marketing has completely solved it, though several have addressed various aspects of the problem. [Montgomery \(1997\)](#) and [Arora, Allenby, and Ginter \(1998\)](#) estimate individual-level response to a marketing mix. However, their models do not determine optimal temporal discounts based on budget constraints. [Tellis and Zufryden \(1995\)](#) develop an optimization model that allows a retailer to determine which brand to discount, by how much, when, and why. However, their model does not allow for customization of discounts to different consumers. [Levedahl \(1986\)](#) develops methods for identifying an optimal coupon of a single face value but does not capture the profit potential that can result from customization of coupons. [Rossi, McCulloch, and Allenby \(1996\)](#) investigate targeted coupons using various information sets available in a database, but their models do not determine optimal temporal discounts based on budget constraints. [Kopalle et al. \(2012\)](#) show how retailers can increase profits by incorporating the heterogeneity in consumer reference prices while deciding pricing policy. However, their model does not optimize the timing of discounts. [Zhang and Krishnamurthi \(2004\)](#) consider future expected consumption patterns (three periods ahead) to determine the optimal depth of discount conditional on a visit. However, they do not consider the optimal timing of discount, which itself can influence consumers' visits.

The current paper extends the literature on price promotion by proposing a model of customized temporal discounts which has four unique characteristics:

- **Response:** A multinomial logit model estimates individual consumer's response to discounts of a target brand. The model allows for heterogeneity among consumers.
- **Timing:** An Erlang-2 timing model estimates individual consumer's inter-purchase time as a function of discounts. This model also allows for heterogeneity across households.

- **Optimization and customization:** A multiple knapsack optimization derives optimal discount values and timing of offers, given either an unconstrained or a constrained budget. The model yields the timing and value of the discount for each household that maximizes the manufacturer's sales or profits.
- **Simulation:** A simulation shows the lift or improvement in profits from adopting customized temporal discounts.

We discuss the customized temporal discounts from the perspective of a retailer. The actual implementation procedure for a brand, with support from manufacturers, would comprise the following procedure: First, determine the optimal discount value and timing for each household. Second, make customized offer known to households as a price-cut or a coupon, at the estimated optimal times through one of various media, such as in-mail flyers, newspaper inserts, emails, mobile texts and pop up or drop down menus in online grocery websites. These steps can be carried out by a promotions intermediary such as Catalina Marketing. Third, apply the discount on purchase of the brand at checkout, with production of appropriate ID for a price-cut or with redemption for a coupon drop. Alternatively, if the shopper is purchasing online apply the discount immediately with the customer's use of an offer number.

The next section describes the models. The third section describes the empirical analysis. The fourth section reports the optimal strategy of timed discounts under different market scenarios. The final section discusses model implementation, limitations and avenues for future research.

Model

This section presents the empirical and optimization models.

Empirical model

We model the consumer's purchase of a product at a specific time as the composite of two events: purchase timing of the product category and brand choice given purchase timing ([Chintagunta 1998; Gupta 1991](#)). The first two subsections describe the choice and timing models, respectively. The third subsection describes the composite decision model.

Brand choice model

Consider a manufacturer who sells brand k at discount d_{ijk} to household i , given incidence j at time t_j . The scale of t_j is fixed by measuring incidence as the time elapsed from the last incidence ($j - 1$). Following [McFadden \(1974\)](#) and [Guadagni and Little \(1983\)](#), we model brand choice given purchase time using a multinomial logit function. Assume that at incidence j of household i , the utility u_{ijk} of brand k at time t_j is:

$$u_{ijk} = \beta_{ik}^0 + \sum_{q=1}^Q \beta_{iq} X_{ijkq} + e_{ijk} \quad (1)$$

where β_{ik}^0 indicates the preference of household i for brand k ; the coefficient vector $\beta_{i1}, \dots, \beta_{iQ}$ captures the household

sensitivity to the Q explanatory variables, X_{ijkq} , such as discount, loyalty, advertising and price; e_{ijk} is the random component of household i 's utility. Without loss of generality we assign X_{ijk1} as the variable for discount d_{ijk} . Assuming the random component of utility, e_{ijk} , is independently distributed as a Gumbel extreme value distribution, the probability of choosing brand k among K competing brands at incidence j and time t_j is given by the familiar multinomial logit model:

$$p_{ij}(Brand = k|\beta_i, X_i) = \frac{\exp(u_{ijk})}{\sum_{l=1}^K \exp(u_{ijl})} \quad (2)$$

where p_{ij} is the probability of household i choosing brand k at time t_j ; to facilitate reading hereafter we write $p_{ij}(k|\cdot) = p_{ij}(Brand = k|\cdot)$. Furthermore, β_i represents the vector $\beta_{i1}^0, \dots, \beta_{iK}^0, \beta_{i1}, \dots, \beta_{iQ}$ and X_i represents the explanatory variables X_{ijkq} .

Timing model

We use the Erlang-2 distribution to model interpurchase time. To incorporate household heterogeneity in purchase times among households we follow Gupta (1991) and Bucklin and Gupta (1992) and use explanatory variables such as average household interpurchase time, inventory levels and category value. Category value reflects the attractiveness of the category and links the purchase incidence to the brand choice decision. Each household's mean time between incidences therefore shifts left or right depending on its general shopping pattern and marketing conditions such as discounts within that category. Formally, for household i the stochastic model for purchase timing at time t_j is given by:

$$p_i(t_j|Y_i) = \alpha_i^2 t_j \exp(-\alpha_i t_j) \quad (3)$$

$$\alpha_i \sim \exp\left(\gamma_i^0 + \sum_{m=1}^M \gamma_{im} Y_{ijm}\right) \quad (4)$$

where $p_i(t_j|Y_i)$ is the probability density of purchase time at time t_j for household i given a set of explanatory variables such as inventory and category value; Y_i represents the vector of M household-specific explanatory variables that may affect the mean of t_j ; α_i = scale parameter of the Erlang-2 distribution for household i , $\gamma_i = \gamma_{i1}^0, \gamma_{i1}, \dots, \gamma_{iM}$ are regression coefficients.

Joint choice and timing model

When a discount d_{ijk} is offered by brand k to the i th household on its j th incidence at time t_j , the joint probability of the household buying the brand k at time t_j , given the household sensitivity to discounts and other explanatory variables can be expressed as:

$$\begin{aligned} p_{ij}(k, t_j|\beta_i, \gamma_i, X_i, Y_i) &= p_{ij}(k|t_j, \beta_i, \gamma_i, X_i, Y_i) p_i(t_j|\beta_i, \gamma_i, X_i, Y_i) \\ &= p_{ij}(k|\beta_i, X_i) p_i(t_j|\gamma_i, Y_i), \end{aligned} \quad (5)$$

where $p_{ij}(k|\beta_i, X_i)$ and $p_i(t_j|\gamma_i, Y_i)$ are respectively given by (2) and (3).

Define the purchase decision of the i th household of brand k at incidence j occurring at time t_j by δ_{ijk} such that

$$\begin{aligned} \delta_{ijk} &= 1 \quad \text{if the household purchases brand } k \\ &= 0 \quad \text{otherwise} \end{aligned}$$

The likelihood function for household i conditional on the discount value, the covariates and the household discount sensitivity is given by:

$$\begin{aligned} L_i &= \int_{R^Q} \int_{R^M} \prod_{j=1}^{J_i} \prod_{k=1}^K p_{ij}(k, t_j|\beta_i, \gamma_i, X_i, Y_i)^{\delta_{ijk}} \varphi(\beta) d\beta \phi(\gamma) d\gamma, \\ L_i &= \int_{R^Q} \int_{R^M} \prod_{j=1}^{J_i} \prod_{k=1}^K p_{ij}(k, t_j|\beta_i, X_i)^{\delta_{ijk}} p_i(t_j|\gamma_i, Y_i) \varphi(\beta) d\beta \phi(\gamma) d\gamma, \quad (6) \\ &= \int_{R^Q} \prod_{j=1}^{J_i} \prod_{k=1}^K p_{ij}(k|\beta_i, X_i)^{\delta_{ijk}} \varphi(\beta) d\beta \int_{R^M} \prod_{j=1}^{J_i} p_i(t_j|\gamma_i, Y_i) \phi(\gamma) d\gamma. \end{aligned}$$

where $\varphi(\beta)$ are the density functions for the parameters β_i with domain of variation R^Q , $\phi(\gamma)$ are the density functions for the parameters γ_i with domain of variation R^M , and J_i denotes the total number of purchases made by household i . Therefore, the overall likelihood is:

$$\begin{aligned} L &= \prod_{i=1}^N L_i = \left[\prod_{i=1}^N \int_{R^Q} \prod_{j=1}^{J_i} \prod_{k=1}^K p_{ij}(k|\beta_i, X_i)^{\delta_{ijk}} \varphi(\beta) d\beta \right] \\ &\quad \times \left[\prod_{i=1}^N \int_{R^M} \prod_{j=1}^{J_i} p_i(t_j|\gamma_i, Y_i) \phi(\gamma) d\gamma \right] \quad (7) \end{aligned}$$

The likelihood structure of the above (7) joint choice and timing model does not provide closed form analytical solutions for the parameter values. So we adopt a hierarchical Bayes framework that allows parameter estimation through simulation (Kopalle et al. 2012). The method involves constructing stationary Markov chains which have the posterior distribution as its stationary distribution. Sequential draws from these distributions are taken to form stable chains which yield parameter estimates to any desired degree of accuracy. We use the built-in Monte Carlo Markov Chain algorithm provided by the Winbugs software. The algorithm simply accepts or rejects each draw based on a pre-specified ratio between the current draw and the previous one. If the ratio exceeds one then the draw is added to the chain else it is rejected and a new one is taken. Parameter estimates are based on these draws. Appendix A provides details.

Optimization model

This subsection develops a constrained optimization model for offering customized temporal discounts in two parts: specification of the objective function and the optimization algorithm.

Specification of objective function

We specify the firm's objective function for profits at the individual level as a generalization of the conventional equation

for calculating net profits at an aggregate level (Blattberg and Neslin 1990, p. 281). This section first derives the aggregate objective function and then extends it to the individual-level. It ends with the intuition about the individual-level objective function.

Aggregate objective function. We assume that the discount offered to a specific household is communicated through one of various media such as mail, email or online. The communication of the discount may incur a cost, which might be positive with traditional media or even 0 with some electronic media. We also assume that processing involves some costs. Furthermore, we consider each discount as a cost relative to the full price a manufacturer could have charged (Tellis 1988). Accordingly, we define net profit $G_N = G - C$, where the gross profit G and total cost C are given by:

$$G = E \times M \times f \tag{8}$$

where E =expected total number of units sold during the promotion period; M =margin per unit when sold at list price; and f =profit share of the manufacturer

$$C = N \times w + E \times r \times (d + a) \tag{9}$$

where w =distribution cost of discount booklet per household; N =total number of households to which discount booklets are sent; d =discount value; a =processing cost and r =redemption rate.

Note that our inclusion of r , $0 \leq r \leq 1$, allows our model to be generalized to a variety of discounts. In the case of an in-store price-cut that all buyers get, r is one. In the case of a coupon, which not every consumer would use, $0 \leq r \leq 1$.

Individual-level profit function. To accommodate household heterogeneity in consumer response to discount and timing, we generalize Eqs. (8) and (9) to form total net profit from individual households:

$$G_N = \sum_{i=1}^N G_i(d, t, T) - C_i(d, t, T). \tag{10}$$

where individual gross profit G_i is given by:

$$G_i(d, t, T) = Mf \left\{ \int_0^t p_i(k, t | \beta_i, D(t) = 0, X'_i, Y_i) dt + \int_t^{t+T} p_i(k, t | \beta_i, D(t) = d, X'_i, Y_i) dt + \int_{t+T}^{\infty} p_i(k, t | \beta_i, D(t) = 0, X'_i, Y_i) dt \right\} \tag{11}$$

where t is the starting time of the offer, T is the length of the period during which the discount offer is valid, and $D(t)$ indicates the discount variable as a function of time t , X'_i are all the remaining choice model variables except discount. The individual cost C_i is given by:

$$C_i(d, t, T) = w + r \times (d + a) \times \int_t^{t+T} p_i(k, t | \beta_i, D(t) = d, X'_i, Y_i) dt \tag{12}$$

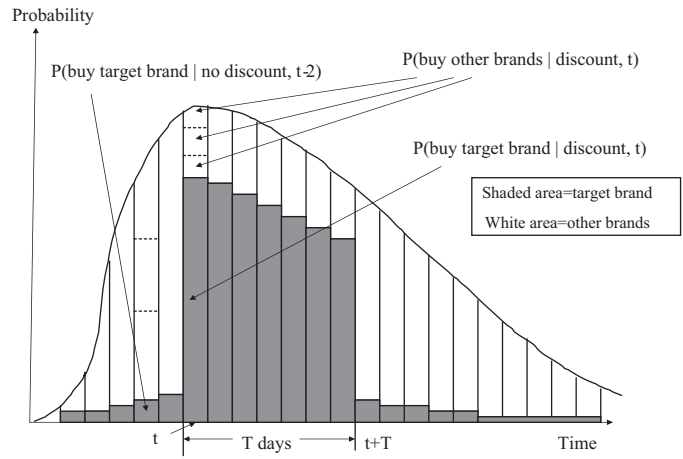


Fig. 1. Joint probability of brand choice and category timing.

The joint probability p_i is given by Eq. (5) but with the subscript j suppressed. We further explain the motivation and the intuition behind Eqs. (11) and (12) in the following subsection.

Intuition about individual-level profit function. A key challenge in determining the optimal temporal discount is to correctly capture the tradeoffs involved in optimizing the value and timing of the discount offer. By jointly using brand choice and purchase timing probabilities, Eqs. (11) and (12) capture the two tradeoffs involved in determining the discount: optimal value and optimal timing.

The intuition of the two tradeoffs is as follows. For discount value, too big a discount to a consumer, who would have bought the brand at a smaller discount, is a lost opportunity for larger margin. Too small a discount to a consumer who would have bought the brand at a lower price is a lost opportunity for one more sale. For discount timing, a discount issued too early may be ignored or lost as discussed in the introduction. A discount issued too late might lead a consumer to buy a rival brand before the consumer gets the offer.

Fig. 1 visualizes these crucial tradeoffs with reference to one manufacturer’s brand, which we call the target brand. The x-axis is time and y-axis is the probability of brand choice given category purchase. The curve in Fig. 1 depicts the estimated Erlang-2 distribution of the interpurchase times of household i . Each column underneath the curve shows the brand choice probabilities for one period (e.g., one day). The shaded portion indicates the probability of household i choosing the target brand at any time. To appreciate the tradeoffs, consider three crude scenarios. If the shaded area is uniformly high, a consumer is very likely to buy the brand, so a discount is unnecessary. If it is uniformly low, then the consumer is unlikely to buy the brand anyway and a discount would probably be futile. If the shaded area is highly uneven, as shown, then the consumer is highly responsive to discounts, so a discount may be appropriate. The optimal timing and value of the discount should maximize the shaded area times the margin per unit less cost of the discount plus distribution and redemption costs.

Algorithm for optimization

The optimization of G_N proceeds to simultaneously maximize individual net profit with respect to discount value, d , time to offer discount, t , and duration of discount T . Denote respectively the value of d , t , and T that maximizes $\pi_i = G_i - C_i$ by d_i^* , t_i^* , T_i^* . The procedure involves a three-dimensional search from the set of plausible values of d , t , and T . If a managerial decision is to offer only discounts that contain values from a finite set at a finite number of time periods, the search is straightforward: select from the set of plausible $\pi_i(d, t, T)$ the largest value $\pi_i(d_i^*, t_i^*, T_i^*)$. Even when d , t and T are continuous-valued, one could discretize the search space of d and t into a finite set of dollar and time values. Discretization is often adopted in practice. For example, each increment in discount value could be 10 cents, and each increment in time could be one day. Consequently, the probability of purchase and time can be evaluated at 10-cent and one-day intervals. Computationally, the discretization scheme takes advantage of efficient vectorized calculations, in which discretized values are stored as vector objects.

To help managers answer the key question – to whom, when and how much to discount – the proposed scheme computes for each household, the optimal discount value, d_i^* , optimal time, t_i^* , and optimal duration, T_i^* . With no budget constraint on the cost of discounting, every household in the entire database would receive a discount of optimal value at an appropriately chosen time. Instead of a complete lack of budget limits, a more realistic scenario is the existence of spending constraints on the promotion campaign. When a budget constraint is present, the problem becomes more complex. It requires constrained optimization in order to solve the set of decision rules.

The constrained problem can be formulated as a resource allocation problem. For computational purposes, discretize the range of discount values $L_d \leq d_u \leq H_d$ into $u = 1, \dots, U$ bins. For the subsequent analyses, set L_d at zero, and H_d at the full price of the brand. Select the step between successive bins d_{u+1} and d_u to be small. Analogously, time t and T are discretized into respectively $v = 1, \dots, V$ and $w = 1, \dots, W$ bins. Furthermore, assume that the total budget for the promotion campaign is B . Let x_{iuvw} denote the decision (1 = yes, 0 = no) as to whether to offer the discount d_u to household i at time t_v with duration T_w .

The problem now can be reformulated as follows:

$$(M1): \text{ maximize } \sum_{i=1}^N \sum_{v=1}^V \sum_{u=1}^U \sum_{w=1}^W \pi_i(d_u, t_v, T_w) \times x_{iuvw} \quad (13)$$

subject to constraints:

$$\sum_{v=1}^V \sum_{u=1}^U \sum_{w=1}^W x_{iuvw} \leq 1 \quad (14)$$

$$\sum_{i=1}^N \sum_{v=1}^V \sum_{u=1}^U \sum_{w=1}^W C_i(d_u, t_v, T_w) \times x_{iuvw} \leq B \quad (15)$$

and

$$x_{iuvw} = 0, 1 \quad \text{for all } i, u, v, w \quad (16)$$

The above formulation results in a type of knapsack problem, such as a hiker choosing the most valuable items to carry in his knapsack, subject to capacity or weight constraints. Metaphorically, (M1) is a multiple-choice knapsack problem in which each of the N households (items) is worth π_i and carries a weight of C_i . The above multiple-choice knapsack problem belongs to a class of decision problems that cannot be solved in polynomial time in the size of the input. The nature of the problem implies that approximation algorithms that cannot guarantee optimality have to be used (Martello and Toth 1990).

A good approximation for the customized temporal discount problem exists and the method of approximation is actually rather intuitive. The method involves the following steps:

- 1 Compute for each household i the respective optimal values of discount, time to offer discount, and discount duration d_i^* , t_i^* , T_i^* , and the corresponding values π_i^* and C_i^* .
- 2 Compute the ratio $V_i^* = \pi_i^*/C_i^*$.
- 3 Rank V_i^* in descending order.

According to the sorted V_i^* sequentially place household i with cost C_i^* into the budget “knapsack” up to the point that the knapsack constraint (15) is violated.

The ratio V_i^* can be viewed as “value per pound of baggage,” or, in the current context, the potential return per dollar invested. Each household is ranked according to its return per dollar, and the best-valued households are included as targets of the planned discount offer.

Empirical analyses

This section presents three applications to test the model: a Japanese application using a price cut, a US application using a coupon and another US application using discounts.

Japanese application: customized price cuts

The next subsections discuss the data, preparation, measures, and results of the Japanese application.

Data

Our data are the daily transactions from a major Japanese chain store. It has approximately 1,300 member stores across Japan and has annual revenues of 70 billion yen. The stores operate under a centralized information system that handles daily transactions from the company’s 2.5 million customers. The chain store also serves as an information and analysis broker for manufacturers. In this application we focus on the laundry detergent market. Four major brands (A, T, S, and R, here to protect anonymity) constitute approximately 95 percent of the market. Brand A has 50 percent of the market. S has the lowest price and market share, while the other two brands are priced approximately the same as A. We apply the proposed model to generate optimal customized temporal price-cuts for the manufacturer of the dominant brand (A).

Preparation

The laundry detergent panel data was collected for the first six months of 1999 from 493 stores that were located in approximately the same region. The extracted data set contains 19,624 purchase transactions of laundry detergents, generated by 6,440 member households. We then selected households that have at least 4 purchases. Unlike large U.S. chain stores, the Japanese stores are typically small neighborhood shops. Shoppers – mostly females who do not drive – often use the same store that is closest to home or work. So we assume store switching is absent. The resulting sample consists of 10,751 transactions from 2,159 households. We set aside the last transaction of each household as a hold-out sample while we used the rest of the data for estimation.

Measures

We derived measures for these variables: LIST PRICE, PRICE CUT, LOYALTY, INVENTORY, INTERPURCHASE TIME and CATEGORY VALUE.

LIST-PRICE is the weekly modal price of the brand sold at the store. In order to capture the behavior of forward looking consumers, we use the concept of future reference price (Rajendran and Tellis 1994; Winer 1986). A reference price is the price paid in the last *n* periods. The future reference price, is the price that the consumer expects on a future time period, based on past prices paid. Thus, it is the consumer’s expectation of prices in the next shopping period. It captures the behavior of forward looking consumers who may compare observed prices with internally stored reference prices to make choice decisions. PRICE CUT is the reference price less transaction price. For a given consumer and brand, reference prices are computed as a function of prices paid by the consumer on previous shopping occasions. Rajendran and Tellis (1994) use the weighted mean

of the previous three prices with the highest weight on the most recent price. Winer’s (1986) uses an ARIMA type model. We follow Winer and fit and ARIMA model for each customer. Transaction price is the actual amount that a household pays at the cash register. LOYALTY signifies brand loyalty and is dynamically evaluated for each individual household. At each transaction, brand loyalty was measured by the share of the brand in the household’s purchases of the category in the 6 months prior to the specific transaction. This *dynamic* measure allows for the evolution of loyalty with the sequence of transactions from the same household (Guadagni and Little 1983). INVENTORY for each household at a point in time was defined as the household’s prior period’s inventory less consumption in the prior period plus any prior period’s purchase. The initial inventory for each household was set at one third of the mean of the household’s total purchases. The consumption for a period was the change in inventory divided by the inter-purchase time. INTERPURCHASE-TIME is the time between two consecutive detergent purchases of a household. CATEGORY VALUE is derived from the choice model for each household at time *t* and is the log of the denominator of the brand choice probability (Bucklin and Lattin 1991).

Estimation and results

This section covers the estimation results of the choice and purchase timing models for the Japanese data. We use the Winbugs software for estimation. See Appendix A for details. Column (2) of Table 1 reports the estimated coefficients. As is customary in Bayesian inference we present the results in terms of the posterior distributions of the parameters and its associated 95 percent probability interval. For both the choice and purchase timing models none of the 95 percent probability intervals include zero. From the mean of discount sensitivity, we

Table 1
 Estimates of the choice and timing models.

(1) Variables	(2) Mean values (detergent data)	(3) Mean values (ketchup data)	(4) Mean values (yogurt data)
Brand choice^a			
LOYALTY	11.89 [11.07 to 12.71]	5.15 [4.66 to 5.63]	3.41 [3.16 to 3.66]
LIST PRICE	-0.10 [-0.11 to -0.09]	-1.08 [-1.19 to -0.97]	-0.56 [-0.64 to -0.48]
PRICE CUT	0.24 [0.15 to 0.32]	0.18 [0.09 to 0.28]	0.14 [0.07 to 0.21]
DISPLAY		1.99 [1.83 to 2.16]	0.80 [0.72 to 0.87]
FEATURE		1.54 [1.40 to 1.67]	0.61 [0.55 to 0.68]
COUPON		0.21 [0.19 to 0.23]	
Timing model			
INVENTORY	0.35 [0.31 to 0.38]	3.17 [2.87 to 3.48]	0.73 [0.30 to 1.16]
INTERPURCHASE TIME	0.045 [0.04 to 0.05]	0.47 [0.41 to 0.54]	0.12 [0.11 to 0.14]
CATEGORY VALUE	-1.17 [-1.33 to -1.01]	-0.30 [-0.34 to -0.25]	-0.17 [-0.19 to -0.15]

^a Brand specific constants are not reported.

see that discounts have significant impact on purchase probability. Despite contextual and cultural differences, the figure seems consistent with price elasticity reported in the marketing literature on U.S. and European consumer goods (Tellis 1988). The significance of the category value coefficient implies that the promotional activities do affect the purchase-timing decision.

US application: customized coupons

The Japanese data did not contain any information on coupons because they were not used as a promotional mechanism in Japan. Another limitation of the Japanese data set was that they came from the same chain store and data from competing stores was not available. To resolve these issues, we conducted a second study with US data. The following subsections describe the data, preparation, measures and results.

Data

Our data are daily household panel data collected by AC Nielsen for the ketchup category, for Sioux Falls, SD from 1985 to 1988. In all there were 2,500 households. Eighty percent of the grocery and drug retail stores in Sioux Falls participated in the study.

Preparation

For convenience in estimating the model we chose the top 4 UPCs by market share, which accounted for roughly 72 percent of the market: Heinz Ketchup, Heinz Ketchup PLS, Hunt's and DelMonte. Their market shares were 35 percent, 17 percent, 13 percent and 7 percent, respectively. We further restricted our data to the 4 largest stores in the market which accounted for roughly 49 percent of the market share. The resultant data contained 10,553 household purchase transactions. We then selected only those households which had at least 5 purchases. The sample size was thus reduced to 9,065 transactions from 970 households. For each household we left out the last observation as a holdout sample.

Measures

We derived measures for LIST-PRICE, PRICE CUT, LOYALTY, INVENTORY, INTERPURCHASE TIME, DISPLAY, FEATURE, coupon value (COUPON), coupon REDEMPTION and CATEGORY VALUE. Except for DISPLAY, FEATURE, COUPON and REDEMPTION all the other measures are the same as the Japanese application and hence will not be described here.

DISPLAY is an indicator variable that shows whether or not there was a display for the purchased brand. FEATURE is an indicator variable that shows whether or not a feature ad was running at the time of purchase. COUPON is the value of the coupon offered to customers. While both store and manufacturer coupon data were available, we used only the manufacturer coupon data because our model is from the perspective of the manufacturer. A check of the data showed that there were no occasions when the manufacturer and the store coupon were redeemed at the same time. Ninety-four percent of the coupons were single coupons while the remaining were double coupons. We constructed a

coupon drop for each store and brand by assuming that if a coupon was redeemed on any given day then the coupon for that brand was available to all customers for that entire week. We then calculated a brand-wise redemption rate for each household. For each household REDEMPTION was calculated by dividing the total number of times the household bought using coupons by the total availability of the coupon for that brand.

Estimation and results

This section discusses the results of the estimation of the choice and purchase timing models for the US data. Column (3) of Table 1 reports the estimated coefficients. The results show coupon face value and price cuts are powerful factors in determining brand choice. Comparing the mean estimated coefficient of sensitivity to price cuts and coupons we see that the mean of coupon sensitivity is higher than that for price cuts. The coefficient for list-price is negative and significant which is what we would expect. The highest coefficient is for brand loyalty, which is also highly significant.

US application: customized discounts

The previous Japanese and US data were for detergent and ketchup respectively and were categories with relatively high storability. We therefore conducted a third study with more recent US data for a perishable item – yogurt. The following subsections describe the data, preparation, measures and results.

Data

We use the IRI Marketing Dataset provided to researchers (see Bronnenberg, Krueger, and Mela 2008, for details). The data contains weekly purchases in multiple product categories by a consumer panel from 2001 to 2006 across US grocery stores. The observations include information on Universal Product Codes (UPCs), prices, discounts, display and feature advertising.

Preparation

For convenience in estimating the model we chose the top 4 brands by market share, which accounted for roughly 73 percent of the market: Yoplait, Danone, Colombo, and Wells. Their market shares were approximately 37 percent, 21 percent, 9 percent and 6 percent, respectively. The data contained 8,319 households with 488,619 purchase transactions. We then selected only those households which had at least 5 purchases. We selected a random sample of 500 households who purchased yogurt in all the years. For each household we left out the last year's transactions as a holdout sample. The sample size was thus reduced to 49,036 transactions from 500 households in the estimation sample and 8,292 transactions in the holdout sample.

Measures

We derived measures for LIST PRICE, PRICE CUT, LOYALTY, INVENTORY, INTERPURCHASE TIME, DISPLAY, FEATURE and CATEGORY VALUE.

Estimation and results

Column (4) of Table 1 reports the estimated coefficients. From the mean estimated coefficients we infer that marketing covariates such as price-cut, display and feature are significant determinants of brand choice. The significance of the category value coefficient also implies that purchase timing is affected by marketing covariates. The highest coefficient is for brand loyalty. The coefficient for list-price is negative and significant which is what we would expect.

Optimal strategy

This section shows how the temporal price-cut and coupon strategy would work for a manufacturer. Using the empirical data and model estimates, we construct some scenarios a typical brand manager would face. In each scenario, we show how optimal temporal discounts compare with those that optimize on value but not timing. We also study temporal discounts on two dimensions: the timing of discounts in comparison to interpurchase times and the profits with respect to loyalty.

The main conclusion that managers can derive from this analysis is that the use of optimal temporal discounts can substantially increase profits. In addition, our simulations suggest some general rules of strategy. We find that optimal discount varies with the size of the margin, the coupon distribution costs, the sensitivity of the market to discounts, and the strength of the brand.

This section has three parts: one for price-cuts based on the Japanese data set, one for coupons based on the US data set and finally one for discounts based on US data. In each case we study a strong brand and a weak brand.

Optimal price cuts

This simulation is on the Japanese data and estimates. We studied the optimal price-cut strategies for the brands with the highest market share (strong brand) and the second lowest market share (weak brand). We did not choose the brand with the lowest share because it had too few transactions.

A price-cut period, T , is defined as the number of days for which price was continuous and constant below the full price in the transactions. To determine T for the Japanese data we first measured the length of T for each store. Next, we took the average of the measured values of T . From the data, the range of the average values of discount periods was 3–51 days, with an overall mean of 6.9 days. Finally, for convenience, we grouped the stores into three categories each with a distinct $T=7, 14,$ and 21 days. We use these values for determining optimal discount periods in the strategic analyses. For competition, we assume that when the customer is offered the price-cut from a target brand, the prices for competing brands are the same as when the household last made a purchase in the category.

The next subsections describe two scenarios: one for the strong brand and another for the weak brand.

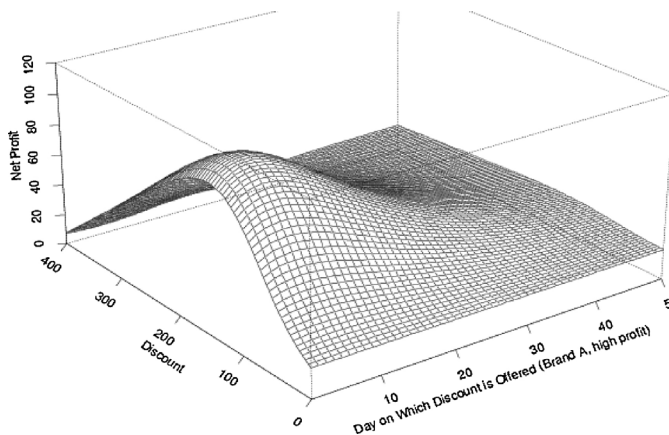


Fig. 2. Net profit for a typical household given value and time of discount.

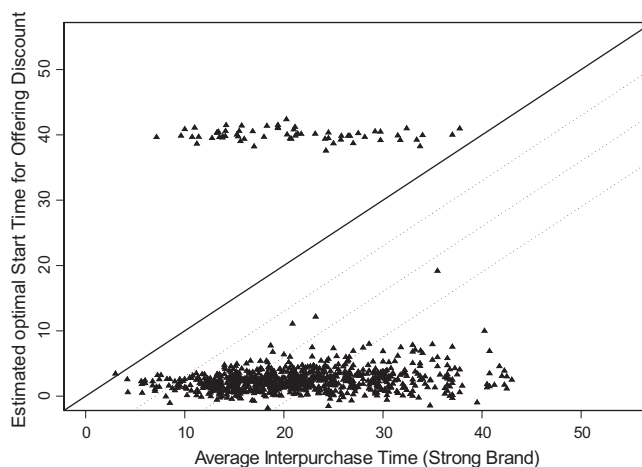


Fig. 3. Optimal time-to-offer versus average interpurchase time (dotted line separated by 7-day intervals).

Strong brand

The first scenario assumes a margin of $M=400$ yen and the manufacturer's share of profit $f=1$.³ This amounts to a margin of approximately 66 percent of the highest list price among all stores. This is a reasonable assumption in normal competitive market environments that feature regular promotions. The distribution cost w was set to 10 yen, under the assumption that there would be cost sharing with other products from the same manufacturer. Because price-cuts are electronically processed and applied at the point of purchase, we assume that the processing cost a is zero and the redemption rate r is one.

Fig. 2 shows the surface of net household profit with respect to price-cuts and time-to-offer for one typical household. For this household, the estimated maximum net profit occurs approximately at a discount of 250 yen and 25 days after the household's last purchase. Fig. 3 shows the optimal time-to-offer plotted against average interpurchase time. The pattern shows that the price-cut should generally be offered to a household a few days ahead of the average household purchase time. The mean time

³ The parameter f allows our model to be applied in contexts where retailer and manufacturers share the cost of discounts (Sethuraman and Tellis 1991).

Table 2
 Model performance from simulation under various scenarios.

Scenario	Average net profit Japanese detergent data (yen)		Average net profit US ketchup data (cents)		Average net profit US yogurt data (cents)	
	Customized temporal	Customized without timing	Customized temporal	Customized without timing	Customized temporal	Customized without timing
Strong brand	108	91	62	49	33	21
Weak brand	40	32	43	39	21	15

to offer t is 5.13 days and the standard deviation is 9.8 days. For most households, the interval $(t, t + T)$ covers the household’s average inter-purchase time. The figure shows that households break into two distinct clusters. The top cluster consists of households whose optimal time is at the boundary of the search space when the termination criterion stops the optimization routine. These are households who should not get discounts because they are likely to continue buying the brand even without discounts. The bottom cluster consists of households whose choice is influenced by discount timing. We compare the customized temporal price-cut to a scheme that customizes value but does not include the timing component. Table 2 summarizes the average net profits derived from the various price-cut schemes. The profit from customized temporal price-cut is 108 yen, while that from customized price-cut without timing is 91 yen. This represents a gain of 18 percent in using customized temporal price-cuts.

Weak brand

We also conducted an analysis for brand S assuming low margin per unit $M=250$ yen and low distribution cost $w=10$ yen. The margin of 250 yen is likely to be an underestimate. Our aim is to use this scenario to study the various “stress” effects of a tight profit margin.

Fig. 4 shows the optimal timing of discounts plotted against the average interpurchase time for the weak brand. The figure shows that for the vast proportion of consumers, the price-cuts should be offered a few days ahead of their next purchase, even when their inter-purchase time is long. As in Fig. 3 we find that

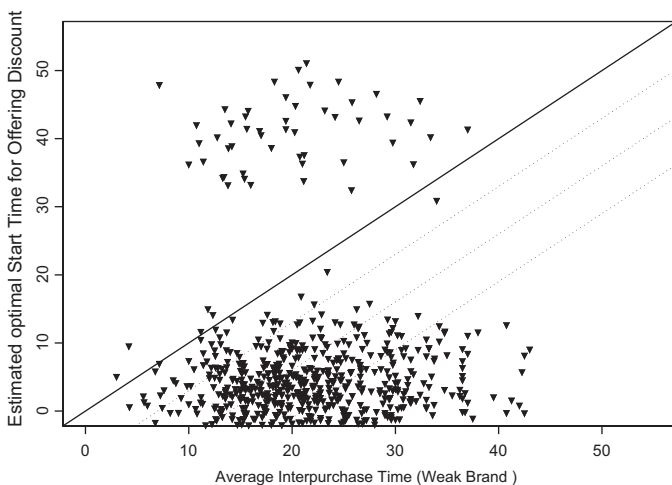


Fig. 4. Optimal time-to-offer versus average interpurchase time (dotted line separated by 7-day intervals).

households cluster into two groups. Again, the customers above the diagonal are those who would buy without discounts. The spread of the data cloud above the diagonal though is larger than in Fig. 3.

Optimal coupons

We performed a similar strategic analysis for the US ketchup data as for the Japanese detergent data. We studied the effects of customized temporal coupons in two cases – the brand with the largest market share (called strong brand) Heinz and the brand with the second lowest share (weak brand) Hunt’s. The parameters of the simulations were set to the following values: margin $M=90$ cents (about 60 percent of maximum list price across all stores), distribution costs $w=5$ cents, expiry of coupon $T, 14$ days. We derived the coupon redemption rate from the actual redemptions in the data. We compared the customized temporal coupons to a baseline case where coupons are customized for value but not time where the coupons are sent out every 7 days. The results are in Table 2.

Strong brand

This scenario shows how customizing the timing of the coupon in addition to its value affects the profits of the manufacturer of the strong brand, Heinz (35 percent market share). Table 2 shows that customizing coupons for both value and time yields a profit of 56 cents. This represents a 22 percent increase in profits over a scheme that is customized only for value but not for time. This result shows that timing coupons can be an extra source of profits for strong brands.

Fig. 5 shows the optimal time-to-offer by average interpurchase time. The pattern shows two distinct segments to which coupons should be sent. A small segment optimally gets the coupon well ahead in their purchase cycle. However, the largest segment should get the coupon just a few days prior to their purchase. This result clearly shows the advantage of our timing model. The strong brand in this dataset benefits less than in the case of the Japanese detergent data. We offer two reasons for this difference. First, the data come from two different countries. It may be that in general Japanese customers are more brand loyal than US customers. Strong brands have more loyal customers. This makes it easier for stronger brands to hold on to their customers and attract switchers with lower discounts than weaker brands. Second, in the US ketchup brands compete on many attribute dimensions while the Japanese detergent brands compete mainly on price.

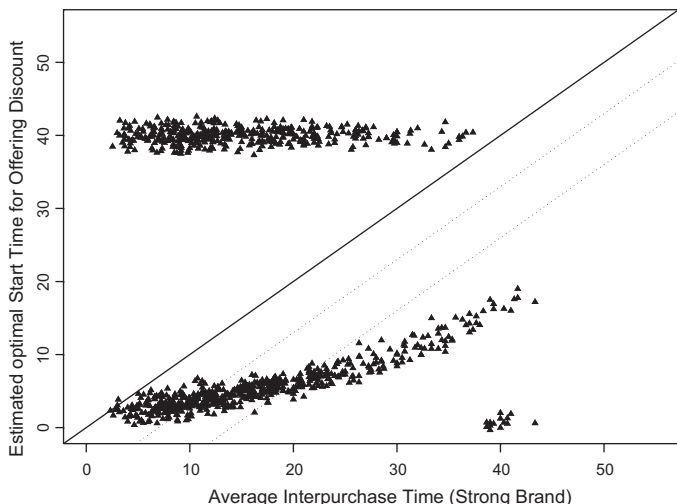


Fig. 5. Optimal time-to-offer versus average interpurchase time (dotted line separated by 7-day intervals).

Weak brand

The weak brand we chose, Hunt’s, had the second lowest market share – about 13 percent. Table 2 shows that customizing coupons for both value and time yields 43 cents of profit. This represents a 23 percent increase in profits over a scheme that entails no customization in time but is customized for value. Thus, this simulation shows that customizing the timing of coupons is more important for a weak brand than it is for strong brands. This may be because the weak brand is usually not in the consideration set of most buyers and a well timed coupon may generate interest in the brand causing people to switch.

Fig. 6 shows the optimal time-to-offer plotted against average interpurchase time. Most consumers get the coupon just a few days prior to their next purchase. In both simulations, the weak brand needs to give the discounts closer to the point of purchase. The reason is that consumers are less likely to be loyal to and thus retain and remember to use coupons of the weak brand, so such firms must deliver coupons as close as possible to the

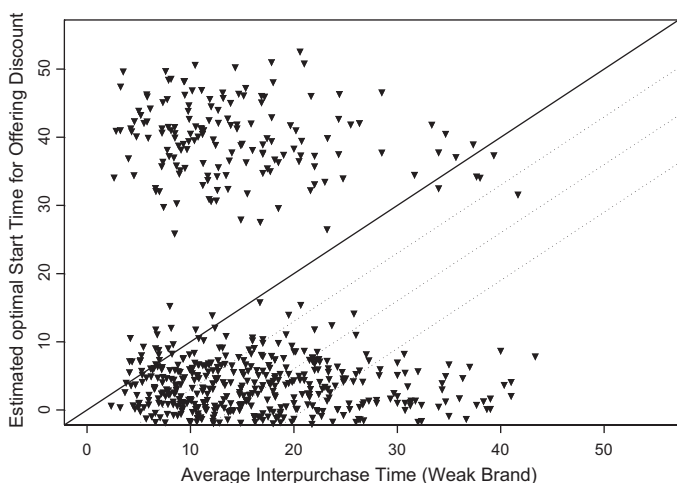


Fig. 6. Optimal time-to-offer versus average interpurchase time (dotted line separated by 7-day intervals).

average interpurchase time. This is distinctly different from the case of the strong brand (see Fig. 5).

Optimal price cuts

We performed a similar strategic analysis for the US yogurt data. We studied the effects of customized discounts for the brand with the largest market share (strong brand) Yoplait and the brand with the second lowest market share (weak brand) Colombo. The parameters of the simulations were set to the following values: margin M at 60 percent of maximum list price across all stores, distribution costs $w = 5$ cents, expiry of discounts T , 14 days. We compared the customized temporal discounts to the baseline case where coupons were customized for value but not time. The results are in Table 2. For the strong yogurt brand we find that customizing discounts for both value and time yields a profit of 28 cents. This represents a 33 percent increase in profits over a scheme that is customized only for value but not for time. For the weak brand we find that customizing discounts for both value and time yields a profit of 21 cents which represents a 40 percent increase in profits over a scheme that customizes only value.

Discussion

While a variety of models for estimating consumers’ response to discounts and optimizing the value of discounts are available, none allow a manufacturer or retailer to optimize both the timing and value of discounts. We developed a model for this purpose, which can be used under a budgetary constraint. The model is general for any type of discount, including price cuts or coupons. The model enables managers to determine to whom, when, and how much to discount a brand, given a budget constraint. More importantly, retailers and manufacturers can conduct specific “what if” analyses for various brands, budgets, and subpopulations of households. Such analyses should allow them to determine the optimal distribution of discount for their firms well before they are distributed.

Our model allows for improvement in profits of 18–40 percent relative to a model that optimizes on value but not timing of discounts. The improvement in profits from using our model may be even more pronounced for those categories where purchase timing is highly responsive to promotions, such as disposable diapers, canned foods, and paper goods. We find that the highest improvement in profits is in the yogurt category perhaps because timing is an important component of the purchase decision dynamics. This lift in profits is important for retailers because of the slim margins they face (Ghemawat 1999).

Our simulation allows for implementation of the models as well as testing “what-if scenarios.” Our limited tests also suggest some strategic patterns which are plausible. We found that profits for the strong brand increases with loyalty probably because loyal customers require less incentive to buy the brand. Weaker brands have to provide larger discounts than stronger brands to entice non-loyal brand switchers. We also found that for most

consumers the optimal time to deliver the discount is ahead of their average interpurchase times.

Validation and robustness

We assess the validity and generalizability of our approach in two ways: performance in hold-out sample and comparison of actual discounts to values from our optimization scheme. We compare the out-of-sample (predictive) performance for the strong brand across our three data sets in terms of mean absolute deviation (MAD). MAD, is a commonly used metric (e.g., Arora, Allenby, and Ginter 1998) which provides an interpretable threshold for model performance. 0.5 implies a predictive choice probability of 0.5, which is equivalent to a coin flip, while 0 implies that choices were predicted with perfect accuracy. We found that the MAD was 0.19, 0.20 and 0.10 for the detergent, ketchup and yogurt data, respectively.

A possible concern with our proposed scheme is the Lucas critique which holds that models based on historical data may not be predictive because the new scheme may induce behavior inconsistent with the past. We address this concern in two ways. First, we compare the actual magnitude of discounts in our data sets to the values derived from our model. For the Japanese detergent data the average price-cut was about 50 yen while the average discount arrived through our model is about 44 yen. For the US ketchup data the average coupon value was about 24 cents and the average coupon value from our model is about 23 cents. For the US yogurt data the average discount value was about 14 cents while the average discount from our model was about 12 cents. Thus in all three data sets the magnitude of actual discounts are similar to the discounts proposed by our model. This implies that consumers will likely react to the proposed discounts in a way consistent with the past. A second way we addressed the Lucas critique was to use a forward looking reference price to calculate our discount variable.

Implementation

This section explains some programs by which our proposal can be implemented. Some of these programs depend on existing promotion instruments while others require new instruments.

Implementation could occur in three ways. First, the most efficient system would be to offer customized discounts electronically by email or mobile text. The growing penetration of Internet and smart phones in the US makes this viable. Online retailers like Peapod (<http://www.peapod.com>) could solve this problem by having the coupon or discount with a promotion code pop up when the consumer logs in. Consumers could then printout the coupon and use it at the retail store of their choice. Alternatively, if they purchase online they can use it immediately by using the promotion code.

Second, retailers could employ the Japanese system. In the Japanese context, the chain is equipped to print out a coupon with the receipt each time the consumer buys at the store. The coupon indicates its value, the product for which it applies, and the start and expiry time when its value is good. All these parameters

come from our model. The model parameters can be updated depending on the computational power available at the store to process the responses to recent purchases. Given the advances in computing this is not unrealistic. In fact the Japanese chain is equipped to do this.

Third, for other consumers, the retailer could print and bundle all coupons for a specific household at specific time periods. The coupons would indicate the value and duration that is optimum for each household. The logistics of this system is the most difficult to implement given the costs needed for customization.

In the context of the Japanese application our model is calibrated for a manufacturer working with a fully cooperative retailer. In this case, whether the chain or the manufacturer carries out the analysis becomes immaterial. The actual implementation depends on who pays for the discounts. If the manufacturer pays for them, then the manufacturer can calibrate its current discounts to match the ones suggested by the model. If the retailer pays for them, then the retailer can calibrate its discounts to match the ones suggested by the model. It can request the manufacturer to pay for or share the costs of any discounts in excess of its current level. The sharing can be accommodated by the parameter f in our model.

The proposed optimization scheme is scalable to the population of households in the market, and can be applied to large data sets. The primary computational overhead is the iterative procedure for computing hierarchical Bayes estimates and search for optimal values. The parameters of the proposed model could be updated at regularly spaced time intervals (such as monthly). Technically, using the parameters as “placeholders” would enable fast and efficient updating of household response information when new data on households are gathered.

While the use of panel data limits the application to panel members several present day contexts permit the calculation of customer response to discounts to the entire market. Examples of such applications include warehouse clubs such COSTCO, credit card companies and cellular phone companies. Similarly, it can be used by firms such as Valucom which offers calling cards at various discounts to consumers. With adjustments, our model can also be used by manufacturers and retailers who send out rebates.

Limitations and future research

Our model has several limitations, which offer opportunities for future research. First, the purchase timing function in our model optimizes the “when to buy” decision and not the “whether to buy” decision. This is because we are trying to optimize the time to send out a discount. Given that most econometric models in marketing treat incidence as a binary decision our optimization could be reformulated to optimize on the binary decision of “whether to buy or not”. Second, the model can be expanded to the context of category management (Hall, Kopalle, and Krishna 2010), instead of the current brand-by-brand approach. However, this would need the incorporation of several category-level factors such as manufacturer incentives across brands and the sales and inventory effects of one

brand on another. Third, our model does not incorporate several characteristics of current purchase such as quantity and store choice. The model could also be refined by incorporating people’s forgetfulness to use coupons, if such data can be collected. Fourth, our model does not incorporate competitive dynamics to account for the situation when all rivals use the model. Fifth, in its current form, the model does not allow for the simultaneous optimization of two decision variables such as coupons and price-cuts. Either the model has to be extended for such a scenario or alternatively, the optimization can be iteratively repeated, each time optimizing for one of the decision variables. Sixth, our current model structure does not allow for complexities such as multiple coupons per customer. Finally, the model also does not allow for an opportunistic retailer who tries to optimize the pass-through of trade deals offered by multiple manufacturers. However, this traditional scenario is becoming less popular in the US, where pay-for-performance schemes reduce such opportunism (Pauwels 2007). Nevertheless, all these remain promising avenues for future research.

Appendix A. Estimation of the hierarchical Bayes model

The hierarchical model described in modeling section is estimated using Markov Chain Monte Carlo (MCMC) methods. For each household (*i*), purchase incidence (*j*) at time *t_j*, and brand (*k*) we have:

A. Conditional choice model: the probability of household *i* for choosing brand *k* is on purchase occasion *j* is given from Eq. (2) as:

$$p_{ijk}(k|\beta_i, t_j, X) = p_{ij}(k|\beta_i, X) = \frac{\exp(u_{ijk})}{\sum_{u=1}^K \exp(u_{iju})} \quad (A.1)$$

where $\beta_i = (\beta_{i1}, \dots, \beta_{iQ})$ are the coefficients for the marketing variables such list-price, loyalty, discounts, coupons, feature advertising and display; *u_{ijk}* is the utility derived from brand *k* and is given as:

$$u_{ijk} = \beta_{ik}^0 + \sum_{q=1}^Q \beta_{iq} X_{ijkq} + e_{ijk} \quad (A.2)$$

Assuming the extreme value distribution for the error *e_{ijk}* leads to the multinomial logit choice model.

Define category value as $\ln \left(\sum_{u=1}^4 \exp(u_{iju}) \right)$.

B. Timing model: The probability of category purchase at incidence *j* is given by the Erlang-2 model specified by Eqs. (3) and (4) as:

$$p_i(t_j|Y) = \alpha_i^2 t \exp(-\alpha_i t) \quad (A.3)$$

$$\alpha_i \sim \exp \left(\gamma_1 + \sum_{q=2}^Q \gamma_q Y_{iq} \right) \quad (A.4)$$

The set of explanatory variables *Y* that determine the probability of purchase includes category value which connects the timing model to the choice model.

C. Priors: The priors for the intercept and *Q* coefficients of the choice model are:

$$\beta_{i,1..k}^0 \sim Normal(\mu_{1..k}^0, \tau_{1..k}^0)$$

$$\beta_{i,1..Q} \sim Normal(\mu_{1..Q}, \tau_{1..Q})$$

where $\mu_{1..k}^0, \mu_{1..Q}$ are the mean and $\tau_{1..k}^0, \tau_{1..Q}$ are the precision values (inverse of variance) of the prior distributions for the coefficients.

The priors for the *Q* coefficients of timing model are: $\lambda_{i,1..Q} \sim Normal(v_{1..Q}, \omega_{1..Q})$ where *v_{1..Q}* are the mean and *ω_{1..Q}* are the precision values (inverse of variance) of the prior distributions for the coefficients.

D. Hyperpriors: We define the hyperpriors for the mean and precision values of choice parameters as follows:

$$\mu_{1..k}^0 \sim Normal(0, 1); \quad \tau_{1..k}^0 \sim Gamma(1, 1)$$

$$\mu_{1..Q} \sim Normal(0, 1); \quad \tau_{1..Q} \sim Gamma(1, 1)$$

Finally, we define the hyperpriors for the mean and precision values of the timing model as:

$$v_{1..Q} \sim Normal(0, 1); \quad \omega_{1..Q} \sim Gamma(1, 1)$$

The above equations are used to specify the model. Estimation is accomplished using MCMC, based on a Gibbs sampling scheme in which we approximate the analytically intractable posterior distribution by sampling from the full conditional distribution. We ran 100,000 iterations using Winbugs software package, combating autocorrelation by thinning the observations – using only every fifth observation. The first 5,000 observations were used for burn-in and the last 15,000 were used for estimation. Thus our parameter estimates for both the timing and the brand choice model were based on the last 15,000 draws. Our results were not sensitive to the priors we chose.

References

Arora, Neeraj, Greg M. Allenby and James L. Ginter (1998), “Hierarchical Bayes Model of Primary and Secondary Demand,” *Marketing Science*, 17 (1), 29–44.
 Blattberg, Robert C. and Scott A. Neslin (1990), *Sales Promotion*, Upper Saddle River, NJ: Prentice Hall.
 Bucklin, Randolph E. and James M. Lattin (1991), “A Two-State Model of Purchase Incidence and Brand Choice,” *Marketing Science*, 10 (1), 24–39.
 Bucklin, Randolph E. and Sunil Gupta (1992), “Brand Choice, Purchase Incidence, and Segmentation: An Integrated Modeling Approach,” *Journal of Marketing Research*, 29 (May), 201–15.
 Bronnenberg, Bart J., Mike Krueger and Carl F. Mela (2008), “The IRI Marketing Dataset,” *Marketing Science*, 27 (4), 745–8.
 Chintagunta, Pradeep (1998), “Inertia and Variety Seeking in a Model of Brand Purchase Timing,” *Marketing Science*, 17 (3), 253–70.
 Ghemawat, Pankaj (1999), *Strategy and the Business Landscape*, Addison-Wesley.
 Guadagni, Peter M. and John D.C. Little (1983), “A Logit Model of Brand Choice Calibrated on Scanner Data,” *Marketing Science*, 2, 208–38.
 Gupta, S. (1991), “Stochastic Models of Interpurchase Time With Time-Dependent Covariates,” *Journal of Marketing Research*, 28 (February), 1–15.

- Hall, Joseph M., Praveen K. Kopalle and Aradhna Krishna (2010), “Retailer Dynamic Pricing and Ordering Decisions: Category Management versus Brand-by-brand Approaches,” *Journal of Retailing*, 86 (2), 172–83.
- Inman, Jeffrey J. and Leigh McAlister (1994), “Do Coupon Expiration Dates Affect Consumer Behavior?,” *Journal of Marketing Research*, 31 (August), 423–8.
- Kopalle, Praveen K., P.K. Kannan, Lin Bao Boldt and Neeraj Arora (2012), “The Impact of Household Level Heterogeneity in Reference Price Effects on Optimal Retailer Pricing Policies,” *Journal of Retailing*, 88, 102–14.
- Levedahl, J. William (1986), “Profit Maximization Pricing of Cents-off Coupons: Promotion or Price Discrimination?,” *Quarterly Journal of Business and Economics*, 25, 56–70.
- Martello, S. and P. Toth (1990), *Knapsack Problems: Algorithms and Computer Implementations*, NY: Wiley.
- McFadden, D. (1974), “Conditional Logit Analysis of Qualitative Choice Behavior,” in *Frontiers in Econometrics*, Zarembka P. ed. New York: Academic Press, Inc, 105–42.
- Montgomery, Alan (1997), “Creating Micro-Marketing Pricing Strategies Using Supermarket Scanner Data,” *Marketing Science*, 16 (3), 15–337.
- Narasimhan, Chakravarthi (1984), “A Price Discrimination Theory of Coupons,” *Marketing Science*, 3, 128–47.
- Pauwels, Koen (2007), “How Retailer and Competitor Decisions Drive the Long-term Effectiveness of Manufacturer Promotions for Fast Moving Consumer Goods,” *Journal of Retailing*, 83 (3), 297–308.
- Rajendran, K.N. and Gerard J. Tellis (1994), “Contextual and Temporal Components of Reference Price,” *Journal of Marketing*, 58 (1), 22–34.
- Rossi, Peter, Robert McCulloch and Greg Allenby (1996), “The Value of Purchase History Data in Target Marketing,” *Marketing Science*, 15, 321–40.
- Salop, Steven and Joseph Stiglitz (1977), “Bargains and Rip-offs: A Model of Monopolistically Competitive Price Dispersion,” *Review of Economic Studies*, 44, 493–510.
- Sethuraman, Raj and Gerard J. Tellis (1991), “An Analysis of the Trade-off between Advertising and Price Discounting,” *Journal of Marketing Research*, 31 (May (2)), 160–74.
- Tellis, Gerard J. (1988), “The Price Sensitivity of Competitive Demand: A Meta-Analysis of Sales Response Models,” *Journal of Marketing Research*, 15 (3), 331–41.
- Tellis, Gerard J. and Fred S. Zufryden (1995), “Tackling the Retailer Decision Maze: Which Brands to Discount, How Much, When, and Why?,” *Marketing Science*, 14, 271–99.
- Varian, Hal R. (1980), “A Model of Sales,” *American Economic Review*, 70 (4), 651–9.
- (2000), “Differential pricing and efficiency,” *First Monday*, 2, <http://www.firstmonday.dk>
- Winer, Russell S. (1986), “A Reference Price Model of Brand Choice for Frequently Purchased Products,” *Journal of Consumer Research*, 13, 250–6.
- Zhang, Jie and Lakshman Krishnamurthi (2004), “Customizing Promotions in Online Stores,” *Marketing Science*, 23, 561–78.