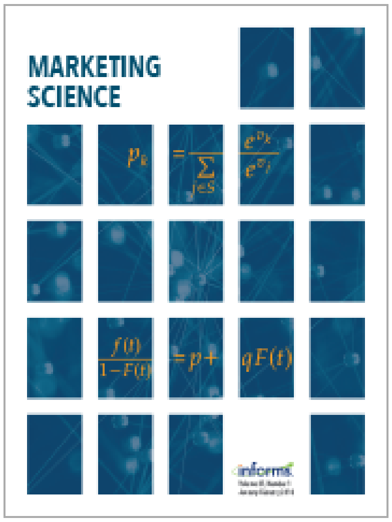


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Price Promotions in Choice Models

John R. Howell, Sanghak Lee, Greg M. Allenby

MARKETING SCIENCE
Vol. 35, No. 2, March–April 2016, pp. 319–334
ISSN 0732-2399 (print) | ISSN 1526-548X (online)

http://dx.doi.org/10.1287/mksc.2015.0948

Price Promotions in Choice Models

John R. Howell
Smeal College of Business, Pennsylvania State University, University Park, Pennsylvania 16802, jrhowell@psu.edu

Sanghak Lee
Tippie College of Business, University of Iowa, Iowa City, Iowa 52242, sanghak-lee@uiowa.edu

Greg M. Allenby
Fisher College of Business, Ohio State University, Columbus, Ohio 43210, allenby.1@osu.edu

Promotions are used in marketing to increase sales and drive profits by temporarily decreasing the price per unit of a good. Some price promotions apply to all quantities (20% off), some have limits on the number of units that can be purchased at a reduced price, and others only offer the discount if the volume purchased is sufficiently high. We develop a model of price promotions in the context of a direct utility model where its effects are incorporated through the budget constraint. Price promotions complicate the estimation and analysis of direct utility models because they induce kinks and points of discontinuity in the budget set. We propose a Bayesian approach to addressing these irregularities and demonstrate the ability of the direct utility model to be used in counterfactual analyses of price promotions. We investigate the stability of utility function estimates for consumers under alternative price promotions, and find that the majority of the effect of a price promotion is through the budget set, not through changes in the utility function. We also investigate the economic value of customized price promotions where the customization includes the value and format of the offer.

Data, as supplemental material, are available at http://dx.doi.org/10.1287/mksc.2015.0948.

Keywords: utility theory; Bayesian estimation; nonlinear pricing; irregular budget sets
History: Received: June 3, 2013; accepted: June 15, 2015; Preyas Desai served as the editor-in-chief and Bart Bronnenberg served as associate editor for this article. Published online in *Articles in Advance* November 20, 2015.

1. Introduction

Firms use price promotions to induce loyal customers to purchase greater quantities and to entice consumers of competitive products to switch brands. Price promotions can be offered uniformly to all respondents, or can be customized in terms of their value and format. Alternative formats include quantity-based discounts, overage charges, and minimum purchase requirements that lead to irregular budget sets with kinks and points of discontinuity that challenge model estimation. An inward kink in the budget set occurs when quantity discounts are offered after a minimum purchase volume occurs. Such promotions are intended to lead some consumers to increase demand quantities to take advantage of the discounted price. An outward kink results when price increases after a certain volume is attained, such as when coupons include a purchase limit. Previous research has documented the economic value of customizing promotional offers to consumers in the form of a simple price reduction for one unit of a good (Rossi et al. 1996). This paper extends the analysis to volume-based discounts under heterogeneous formats.

A challenge in analyzing different promotional formats is addressing many different price schedules. One approach is to use dummy variables to represent each promotional type. This approach is commonly used in choice models when display and feature variables are included in the analysis (see, e.g., Guadagni and Little 1983, Villas-Boas and Winer 1999, Mehta et al. 2010), and could be extended to include other types of promotional variables. However, as the number of products and promotional offerings increase, this approach requires a large number of dummy variables to investigate each different type of promotion. The dummy variable approach also limits the types of counterfactual analyses that can be considered to those that conform to the experimental conditions.

In addition, dummy variables cannot fully account for the quantity-based pricing. The challenge here is that the final price charged depends on the quantity purchased. This leads to a circular dependency in the demand equations as the observed demand depends on price and price depends on observed demand. We develop a direct utility model that allows for nonlinear pricing. This direct utility approach separates the price from the utility function and resolves the circular dependency. This allows the full pricing schedule to be represented in the model rather than relying on simplifying assumptions. One challenge of nonlinear pricing is that the budget frontier is no longer linear or fully differentiable. The kinks induced at the price changes necessitate a new method for solving the utility maximization problem and fitting the stochastic model. The framing of the demand function to fully account for the nonlinear price schedule and the method for estimating the model are the primary contributions of this paper.

Our structural model can be used to analyze quantity-limited discounts, volume discounts, and discount schedules offering multiple discount points. In addition, it nests the constant price discount model of Kim et al. (2002). The model separates a pure price effect of a promotion on the budget set from its effect on the utility function. The model also accommodates full heterogeneity in consumer responses to promotions and allows for a unique pricing schedule for each consumer. This enables the model to consider the mass customization of prices and formats, and facilitates analysis of how different pricing schedules impact distinct consumers.

We demonstrate our model using a conjoint analysis of carbonated soft drinks. The conjoint analysis was designed to allow consumers to specify an unrestricted volume for each choice offering for a variety of price and promotional conditions. We find that price promotions have little effect on the utility function once the pure price effects of the budget set are modeled. We investigate the benefit of individual price customization in discount types and discount levels, and show that consumers respond to various types of promotions differently. We also investigate how the promotion types can be used to target specific consumers.

The remainder of the paper is organized as follows. In §2 we develop a general model of quantity-based discounts and discuss challenges associated with kinks and points of discontinuity in the budget set. Estimation algorithms are developed that do not rely on first-order conditions (FOCs) typically used for continuous models of demand. In §3 we discuss the conjoint analysis used to demonstrate our model. Section 4 investigates potential uses of our model in offering heterogeneous promotional formats and price discounts to consumers. We find that different promotion types have a differential effect on the source of volume among consumers. We conclude in §5 by discussing model limitations and possible extensions.

2. A Direct Utility Model for Price Promotions

Quantity-based pricing and price promotions have received significant attention in the marketing and economics literature. The economics literature has primarily focused on the theoretical effect of quantity discounts and their use as a price discrimination mechanism. We extend this literature to develop a general approach to estimating choice models with price promotions, to investigate the empirical effects of price promotions on demand, and to identify the benefits of price promotion customization that extend beyond a simple price discount.

We develop a direct utility model for price promotions that allows for corner and interior solutions while distinguishing the influence of the price promotion on the budget set versus the utility function. An alternative approach to using a direct utility model is use of an indirect utility function that specifies demand, x, as a function of product attributes (a), prices (p), and expenditure (E) or $x = f(a, p, E)$. A limitation in using an indirect utility function is that the budget and utility effects can be difficult to separate, especially when a flexible functional form is assumed. In a general price promotion, the price depends on the quantity purchased. This leads to an indirect utility function that is implicitly instead of explicitly defined

$$x = f(a, p(x), E).$$

Quantity-dependent prices complicate analysis in a number of other ways as well. First, Roy's identity cannot be used to justify the form of the demand function as arising from a constrained maximized utility problem because doing so requires exogenous and constant prices (see Deaton and Muellbauer 1980). The discontinuities induced by the price schedule rule out the use of FOCs for estimation and analysis. Second, there may be elements of the demand vector x equal to zero where consumers are observed to not purchase any of a choice alternative. This occurs often in marketing data sets where analysis is conducted at a disaggregate level. It is unclear what price should be used (e.g., the base price, the average price, the marginal price) when a consumer is observed to not pay any price. Third, we show below that the general price schedule generates kinks and discontinuities in the budget set that result in a piling up of demand at specific points. Any coherent analysis of demand data associated with nonconstant prices must account for this possibility.

To address the discontinuities and kinks in the budget set, we use a direct utility model that does not rely on FOCs for parameter estimation and analysis. Existing empirical work involving price schedules has only allowed for corner solutions where at most one of the choice alternatives is selected. Allenby et al. (2004) propose a model with quantity price discounts useful when only one brand is selected. Iyengar and Jedidi (2012) propose a discrete choice model where quantity bundles are treated as distinct choice alternatives. It is not possible to adapt these models to the situation where multiple alternatives can be chosen, as is present in our data set. In addition, these papers rely on prices monotonically decreasing with quantity, while our model handles the more general case with prices increasing or decreasing.

The direct utility model was originally proposed for empirical work by Wales and Woodland (1983) and has since been adapted to a number of different marketing situations (see Kim et al. 2002, Bhat 2005, Satomura et al. 2011b, Lee et al. 2013, Howell and Allenby 2014). We extend this line of literature to handle quantity-based pricing. We begin with a simple two-good demonstration that allows us to graphically show the solution procedure. We then demonstrate a basic model that incorporates a single kink point into the expenditure constraint assuming continuous demand. Multiple kink points are possible and a straightforward extension of the model. We then show how to modify the model to account for the fact that customers can only purchase discrete quantities using the procedure described in Lee and Allenby (2014) that does not require the use of FOCs.

2.1. Graphical Demonstration

Figure 1 is a graphic depiction of the utility maximizing solution with two goods, i.e., a focal inside good and an outside good. In this example, the price of the inside good is reduced by 25% for all quantities purchased. Because the price discount simply shifts the budget constraint, the standard model of direct utility maximization based on Kuhn-Tucker conditions is adequate. The point of tangency between the linear budget constraint and the indifference curve represents the consumer's optimal purchase quantity.

Figures 2 and 3 illustrate the utility maximizing solution to more complicated price promotions. Figure 2 is a price promotion that is limited to a fixed purchase quantity. The budget constraint bends outward until the purchase limit is reached. It then becomes parallel to the original budget constraint for quantities of the inside good beyond the limit. This leads to a kink in the budget constraint at the discount limit, which is set equal to two in this example. With this nonlinear budget constraint, there are many possible indifference curves that are optimal and pass through the kink point. This leads to a large number of observations at the kink point. It is also no longer possible to use the Kuhn-Tucker tangency conditions to find the utility maximizing solution.

Figure 3 is an example of a traditional volume discount where purchase quantities greater than some lower limit are sold at a discount. The original budget constraint is matched until the quantity threshold is reached and then rotates outward for higher purchase quantities. The lower limit to the promotion produces an inward kink to the budget set. This inward kink results in reduced purchases near the kink point as consumers increase purchases to receive the discount. Even though there is no mass build-up of purchases at the kink point, the tangency condition of the Kuhn-Tucker solution is not sufficient to ensure optimality.

Figure 2 (Color online) Example of Quantity Limited Discount

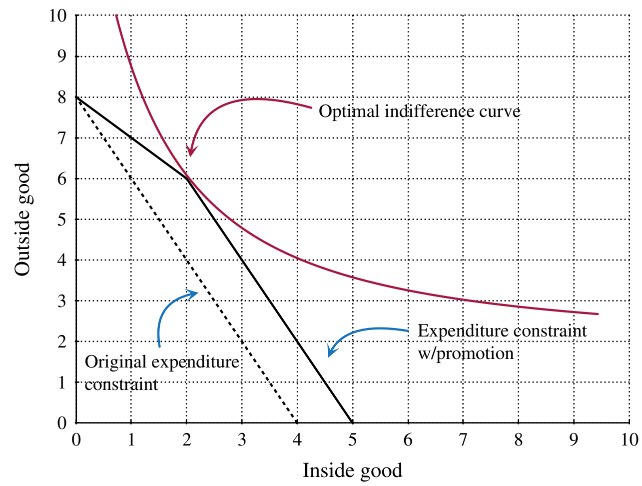


Figure 1 (Color online) Example of Fixed Percent-Off Promotion

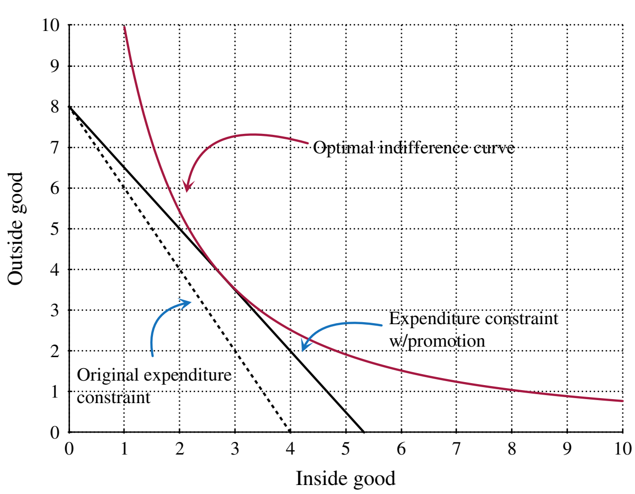
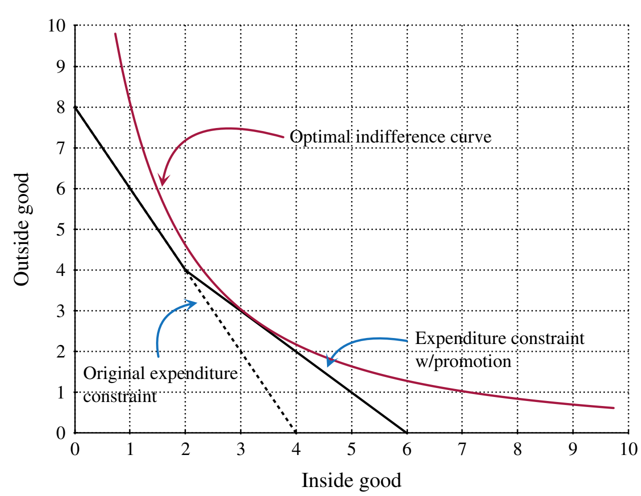


Figure 3 (Color online) Example of a Volume Discount



2.2. A General Model

We expand the model to handle an arbitrary number of products with a mixture of pricing schedules where consumers are assumed to maximize utility subject to a budget constraint. For simplicity of notation, we focus on a single consumer's utility maximization problem and suppress the respondent specific subscript. We can write the observed demand as a solution to the following constrained optimization problem where x^*, z^* are the observed optimal demands for the inside and outside goods respectively

$$(x^*, z^*) = \underset{x,z}{\arg\max}\; U(x,z) \quad \text{s.t.} \quad \sum_{i=1}^{I} [p_{li}\min(x_i, \tau_i) + p_{hi}\max(0, x_i - \tau_i)] + z = E, \tag{1}$$

where τ_i is the quantity limit at which the discount or surcharge is incurred, $p_{h\cdot}$ is the price for each good x_i above the threshold τ, and $p_{l\cdot}$ is the price below the threshold.[1] The price of the outside good, z, is normalized to 1. This specification leads to the total price paid for quantities above the threshold τ_i equal to $p_{li}\tau_i + p_{hi}(x_i - \tau_i)$ for good i.[2]

The presence of one or more kink points $\{\tau_i\}$ invalidates use of the Kuhn-Tucker conditions (see Mas-Colell et al. 1995) for identifying optimality conditions associated with observed demand. The expenditure set is no longer convex for the case of quantity discounts, which further complicates the analysis. An efficient solution to the utility maximization problem is to divide the expenditure constraint into a set of regions composed of linear expenditure constraints and boundary conditions. Provided the partitioning is exhaustive, each partition can be maximized separately and then compared to determine the overall optimum. With only two inside goods, this procedure would correspond to four unique partitions. In general, a single kink point leads to 2^I partitions where I is the number of inside goods with positive τ.

The partitioned utility maximization problem for two inside goods becomes

$$\mathbb{P}_1: \underset{x_1,x_2,z}{\arg\max}\; U(x_1,x_2,z) \quad \text{s.t.} \quad p_{l1}x_1 + p_{l2}x_2 + z = E, \quad 0 \le x_1 \le \tau_1,\; 0 \le x_2 \le \tau_2, \tag{2}$$

$$\mathbb{P}_2: \underset{x_1,x_2,z}{\arg\max}\; U(x_1,x_2,z) \quad \text{s.t.} \quad p_{l1}\tau_1 + p_{h1}(x_1-\tau_1) + p_{l2}x_2 + z = E, \quad \tau_1 < x_1,\; 0 \le x_2 \le \tau_2, \tag{3}$$

$$\mathbb{P}_3: \underset{x_1,x_2,z}{\arg\max}\; U(x_1,x_2,z) \quad \text{s.t.} \quad p_{l2}\tau_2 + p_{l1}x_1 + p_{h2}(x_2-\tau_2) + z = E, \quad 0 \le x_1 \le \tau_1,\; \tau_2 < x_2, \tag{4}$$

$$\mathbb{P}_4: \underset{x_1,x_2,z}{\arg\max}\; U(x_1,x_2,z) \quad \text{s.t.} \quad p_{l1}\tau_1 + p_{l2}\tau_2 + p_{h1}(x_1-\tau_1) + p_{h2}(x_2-\tau_2) + z = E, \quad \tau_1 < x_1,\; \tau_2 < x_2. \tag{5}$$

Each of the partitions involves a linear constraint and up to two boundary conditions and thus can easily be solved using the Kuhn-Tucker conditions as described in Howell and Allenby (2014). Once the optimal solution for each partition is found, the demand corresponding to the highest utility partition represents the optimal demand.

Note that this technique can handle an arbitrary number of kink points as well as other types of discontinuities. The challenge becomes properly defining the pricing structure and budget frontier that the consumer faces and dividing that space into an exhaustive set of linear constraints that can be individually maximized and then compared. While the complexity of the solution increases exponentially as kink points are added, the basic structure of the problem does not change. In practice, most pricing schedules involve a relatively small number of price changes.

2.3. Likelihood

Error terms are introduced into the model specification to account for unobserved factors and to derive the model likelihood for empirical estimation. The model framework is general and can be applied to a variety of utility functions. We demonstrate the model with a variation of the linear expenditure system that has been used in marketing (see Satomura et al. 2011a, b). The model allows for corner and interior solutions as well as cross-price effects due to an income effect. The utility function takes the form

$$U(x,z) = \sum_{i=1}^{I} \psi_i \log(\gamma x_i + 1) + \psi_z \log(z), \tag{6}$$

where ψ_i captures the latent product specific preference, γ allows for individual specific rates of satiation, x_i is the demand for the inside products, and z is the outside good. For the utility function to be valid, ψ_i must be positive. Thus we reparametrize it such

[1] In general argmax does not guarantee a unique solution. With a proper utility function, however, it is expected that the solution will be unique.

[2] This specification only applies when the price schedule involves discrete price breaks. If the price schedule is continuous, a piecewise approximation would be required or the method would not apply. We have not observed a situation with a continuous price schedule in practice.

that $\log(\psi_i) = V_i + \varepsilon_i$ where V_i is the deterministic portion of the preference parameter and ε_i is the stochastic portion. As is common in random utility models, we parametrize V_i as a function of product attributes and part-worth utility weights such that $V_i = a_i' \cdot \beta$ (McFadden 1986). The a_i variables represent the observed independent variables that contribute to the total utility for the good. These variables could be product attributes such as brands, package sizes, and features or contextually relevant variables such as a dummy indicating that a product is on promotion. The β is a vector of regression weights that are shared across all products, but are unique to the individual. The error term, ε_i, captures the influences on product utility not accounted for in the deterministic portion of the utility function. This could include omitted variables or errors in optimization on the part of the individual.

The full set of ψ parameters are not theoretically identified as multiplication by any positive constant does not change the utility maximizing solution (Deaton and Muellbauer 1980). Thus, for empirical identification purposes, we divide the utility function by ψ_z and relabel ψ_i/ψ_z as ψ_i.

The full utility specification is

$$(x^*, z^*) = \underset{x,z}{\arg\max} \sum_{i=1}^{I} \psi_i \log(\gamma x_i + 1) + \log(z) \tag{7}$$

$$\text{s.t.} \quad \sum_{i=1}^{I} \big[p_{li} \min(x_i, \tau_i) + p_{hi} \max(0, x_i - \tau_i) \big] + z = E,$$
$$x \geq 0, z > 0,$$

$$\text{with} \quad \psi_i = \exp(a_i \beta + \varepsilon_i). \tag{8}$$

The econometric challenge is to estimate the probability of the parameters given the observed data. We take p_l, p_h, and τ as exogenously determined and estimate the remaining unknown parameters β, γ, and the quantity of the outside good, z, or the budget constraint, E.

Because of the adding-up condition in the expenditure constraint, $E = \sum_{i=1}^{I} p_i(x_i) + z$ where $p_i(\cdot)$ is the price function for good i, it is not possible to separately consider the outside good, z, and the expenditure constraint, E (Deaton and Muellbauer 1980). If we could observe the quantity of the outside good purchased or the total expenditure constraint, we could infer the other. The outside good can be thought of as a composite good that competes with the inside goods for budget allocation or as money saved that can then be spent in other product categories. We follow Allenby et al. (2004) and estimate E.

The parameters of the utility function are identified by the changes in volume observed across the choice tasks. Specifically, the β parameters are identified by the variation in demand across the different choice offerings. The γ parameter is identified by the variation in total amount purchased for each offering and measures the rate at which customers satiate on each product. The expenditure constraint E is identified due to the changes in total quantity purchased and amount of money spent. This captures the substitution rate between the inside and outside goods as the prices of the inside goods change.

We now demonstrate this solution for the previously discussed two inside good model that has four partitions. Consider the case wherein the observed demand falls in the first region. This region is defined by the constraint

$$\begin{gathered} p_{l1}x_1 + p_{l2}x_2 + z = E, \\ 0 \leq x_1 \leq \tau_1, 0 \leq x_2 \leq \tau_2. \end{gathered} \tag{9}$$

The important consideration in this partition is that the demand is constrained by an upper and lower boundary.

The first step is to form the auxiliary function using the Kuhn-Tucker conditions and compute the FOCs. The general FOCs are

$$\begin{aligned}
&\frac{\psi_i \gamma}{\gamma x_i + 1} - \frac{p_{li}}{z} \leq 0 && \text{if } x_i = 0, \\
&\frac{\psi_i \gamma}{\gamma x_i + 1} - \frac{p_{li}}{z} = 0 && \text{if } 0 < x_i < \tau_i, \\
&\frac{\psi_i \gamma}{\gamma x_i + 1} - \frac{p_{hi}}{z} \leq 0 \leq \frac{\psi_i \gamma}{\gamma x_i + 1} - \frac{p_{li}}{z} && \text{if } x_i = \tau_i, \\
&\frac{\psi_i \gamma}{\gamma x_i + 1} - \frac{p_{hi}}{z} = 0 && \text{if } x_i > \tau_i,
\end{aligned} \tag{10}$$

$$\text{where } z = E - \sum_{i=1}^{I} \big[p_{li} \min(x_i, \tau_i) + p_{hi} \max(0, x_i - \tau_i) \big].$$

Solving the FOCs for ε_i leads to

$$\begin{aligned}
&\varepsilon_i < g_{li} && \text{if } x_i = 0, \\
&\varepsilon_i = g_{li} && \text{if } 0 < x_i < \tau_i, \\
&g_{li} < \varepsilon_i < g_{hi} && \text{if } x_i = \tau_i, \\
&\varepsilon_i = g_{hi} && \text{if } x_i > \tau_i,
\end{aligned} \tag{11}$$

where

$$\begin{aligned}
g_{li} &\equiv -a_i'\beta + \log \frac{(\gamma x_i + 1) p_{li}}{\gamma z}, \\
g_{hi} &\equiv -a_i'\beta + \log \frac{(\gamma x_i + 1) p_{hi}}{\gamma z}.
\end{aligned} \tag{12}$$

These conditions correspond to a mix of density and mass contributions to the likelihood. The mass contributions occur because a number of different values of ε_i would lead to the same observed outcome as the optimal demand occurs at a boundary condition. Thus we need to integrate across the possible values of the error term.

We partition the elements of $\mathbf{x}$ into the following sets: The set $A = \{i\colon x_i = 0\}$, the set $B = \{i\colon 0 < x_i < \tau_i\}$, the set $C = \{i\colon x_i = \tau_i\}$, and the set $D = \{i\colon x_i > \tau_i\}$. The likelihood is therefore

$$\begin{aligned}\Pr(\mathbf{x_i}) &= \Pr(x_A = 0, 0 < x_B < \tau_i, x_C = \tau_i, x_D > \tau_i)\\ &= |J_{B\cup D}| \Pr(\varepsilon_A < g_{lA}, \varepsilon_B = g_{lB}, g_{lC} < \varepsilon_C < g_{hC}, \varepsilon_D = g_{hD})\\ &= |J_{B\cup D}| \times F(g_{lA}) \times f(g_{lB})\\ &\quad \times (F(g_{hC}) - F(g_{lC})) \times f(g_{h,D}),\end{aligned} \tag{13}$$

where f is the Probability Density Function (PDF) of the error distribution and F is the Cumulative Distribution Function (CDF) of that distribution. The Jacobian of $\varepsilon_{B\cup D}$, $J_{B\cup D}$, is defined as

$$J_{ij} = \frac{\partial g_{\cdot i}}{\partial x_j} = \frac{1}{x_i + 1} I(i = j) + \frac{p_{\cdot j}}{z}, \tag{14}$$

with $\cdot = l$ if $i \in B$ and $\cdot = h$ if $i \in D$.

The model can be extended to handle multiple kink points as is the case with buy-one-get-one (BOGO) promotions. Each additional price point introduces an additional regime that must be considered. The challenge is tracking the appropriate prices and expenditure constraint for the chosen regime.

2.4. Discrete Demand Estimation

Our proposed model in (7) and (8) assumes that consumers purchase a continuous quantity, when in reality they are often restricted to purchasing whole numbers of packages. We accommodate this indivisibility of demand by introducing an integer constraint for the quantities of the inside goods. We briefly summarize this method here; it is described in detail in Lee and Allenby (2014). The utility function is modified as follows

$$\begin{aligned}(x^*, z^*) = \underset{x,z}{\arg\max} &\sum_{i=1}^{I} \psi_i \log(\gamma x_i + 1) + \log(z)\\ \text{s.t.} &\sum_{i=1}^{I} \big[p_{li} \min(x_i, \tau_i)\\ &\qquad + p_{hi} \max(0, x_i - \tau_i)\big] + z = E\\ &x \in \{0, 1, 2, \ldots\}, z > 0.\end{aligned} \tag{15}$$

The addition of the integer constraint modifies the likelihood function so that each observation can be rationalized by a set of error realizations. Instead of associating FOCs with observed demand, our proposed estimator relies on directly comparing the feasible integer points using an integer-programming like algorithm.

We leverage properties of our model structure to significantly reduce the computational burden of estimation without use of FOCs. The first property is that the utility function is concave. The second property is that it is possible to find the implied utility maximizing solution in the continuous case, which identifies the neighborhood in which the constrained maximization occurs. These two aspects of our model enable us to greatly reduce the set of solutions we need to consider for model estimation.

Following §§5.1–5.3 of Lee and Allenby (2014) we use the error augmentation procedure to solve the stochastic maximization problem. For a given draw of the set of parameters and the error term, we systematically compare various solutions to the integer constrained utility maximization problem to determine whether the draws are consistent with the observed data. This allows us to integrate across all possible realizations of the error terms to estimate the underlying parameter values. Details of the estimation algorithm are provided in Appendix A.

3. Empirical Analysis

We investigate the performance of our model with a conjoint data set of soft drink purchases. While analysis of scanner panel data is possible with the model, currently available data sets do not record the necessary information to fully characterize the price schedule a respondent faces. The variables reported in syndicated data sets include the total price paid for a good and dummy variable indicators for feature advertising and the presence of various displays. We note that point of sale systems could technically be programmed to record the full price schedule for respondents and we hope that our research will encourage more detailed reporting of the purchase environment.

We used a three-cell within-subject design for data collection. Each cell corresponds to a common promotional activity. The first cell uses a simple percentage discount promotion that is not restricted by quantity. The second cell corresponded to a limited quantity discount that only discounted the first three units leading to an outward kink in the expenditure set. The third cell received a volume discount where the price decreases for each unit purchase beyond the first two, leading to an inward kink in the expenditure set. Survey Sampling International (SSI), a commercial panel company, fielded the study. SSI specializes in contacting respondents for marketing research questionnaires and is widely used in commercial marketing research. The sample was chosen to be representative of consumers of soft drinks in the United States. Screening questions were used to ensure that respondents regularly consumed the soft drink brands included in

Table 1 Product Features

Brands	Package sizes	Price
Coca-Cola Classic	12-pack 12 oz. cans	See Table 2
Diet Coke	6-pack 24 oz. bottles	
Pepsi	2-liter bottle	
Diet Pepsi		
Dr Pepper		
Diet Dr Pepper		
Mountain Dew		
Diet Mountain Dew		
Sierra Mist		
Diet Sierra Mist		
7-Up		
Diet 7-Up		
Sprite		
Sprite Zero		
Orange Fanta		
Orange Crush		

Table 2 Prices Customized by Package Size

12 oz. cans ($)	24 oz. bottles ($)	2-liter bottle ($)
3.49	2.99	1.29
3.69	3.19	1.39
3.85	3.35	1.49
3.99	3.49	1.59
4.15	3.65	1.69
4.29	3.79	1.79
4.49	3.99	1.89

our study. The attributes and levels used in the conjoint task are presented in Tables 1 and 2.

Sixteen soft drink brands were included in the study with national distribution across three manufacturers. These represent the 10 most popular brands in the United States market as measured by market shares and six additional competitors (Beverage Digest 2012). The total market share of the included brands exceeds 65% of the United States carbonated soft drink market. In addition to the brand name, we varied the package sizes to include the most commonly available packages sold at grocery stores. Pricing was consistent with national prices and customized for each package size, but varied randomly across the brands.

Each respondent answered a total of 20 conjoint tasks; 10 tasks without a promotion and 10 with a promotion. The 10 treatment tasks were randomly interspersed with the control tasks to minimize any order effects. An example screen from the conjoint study is presented in Figure 4. Example stimuli for the remaining cells are presented in Appendix B.

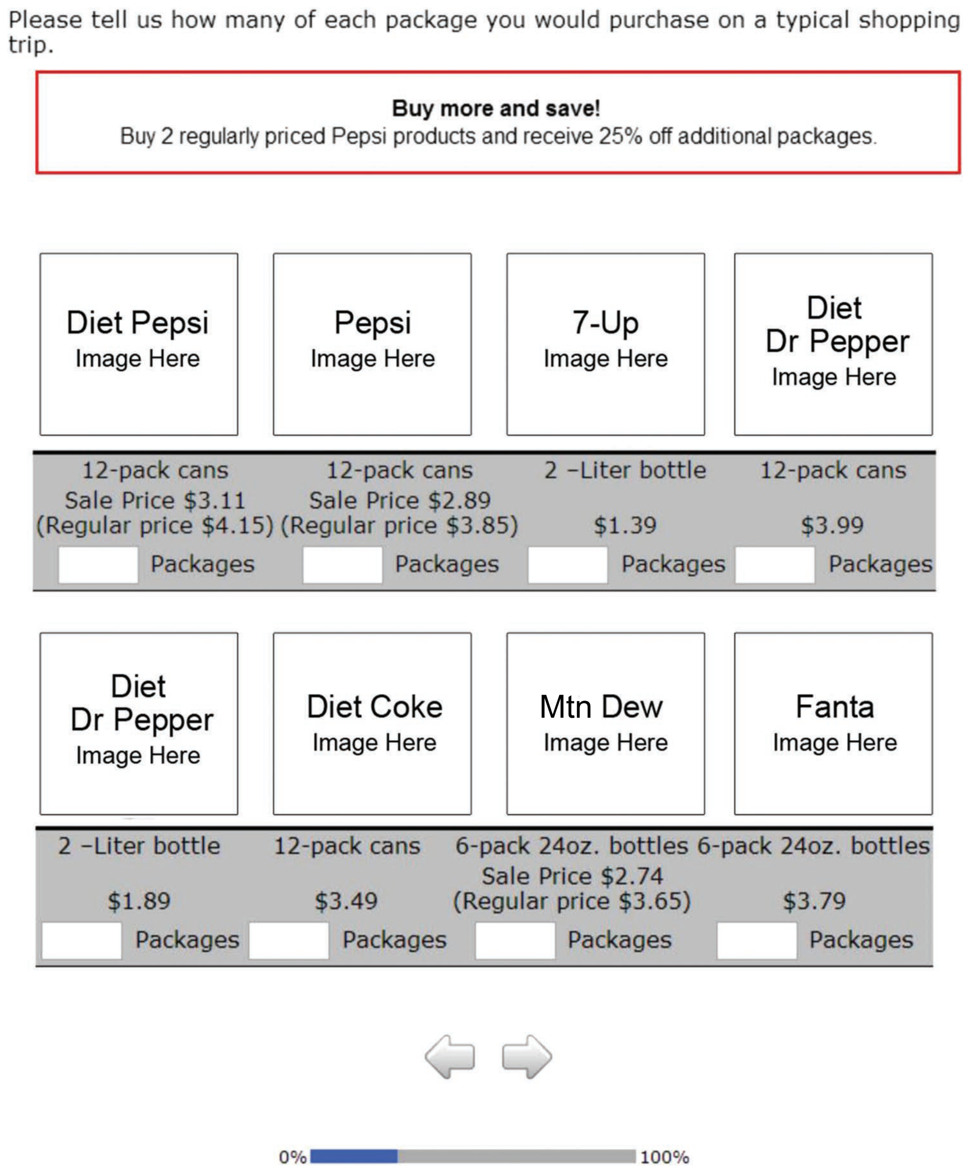
Before estimation the data was cleaned to remove respondents who did not properly follow the survey instructions or did not allow for estimation. Of the 790 respondents, 43 answered zero for all offerings in all choice tasks or did not complete the choice exercise. This could indicate that respondents did not properly consider the survey exercise and screening questions or were exclusive purchasers of a brand not included in the survey instrument. The proportion of nonpurchasers is consistent with other conjoint analysis surveys that we have seen.

An additional 54 respondents were removed because they answered positive quantities for the majority of product offerings regardless of price or package sizes thus indicating unrealistic response patterns. The cut-off point for inclusion of the survey was determined based on the number of zero responses out of 160 provided by each respondent. To be removed, the respondents had to answer positive demands for half of the choice tasks. These respondents show clear evidence of not treating the survey instrument seriously and thus indicate unreliable responses. Four additional respondents were removed for extreme response patterns. In addition to those respondents removed for data quality problems, 10 had responses that could not be rationalized under the benchmark model although they could be estimated under the volume discount model. To make the comparison consistent, these 10 respondents were removed from the final analysis. The final sample size for our study is 679 respondents with 249 in the flat percent off cell, 233 in the limited quantity cell, and 197 in the volume discount cell. The conjoint data show that a zero is the most frequently observed quantity (86.5%), and there exist a significant number (33.0%) of conjoint tasks in which none of the choice alternatives are chosen. An average respondent spends $5.05 for 1.88 units of the available items that are composed of 1.08 different alternatives. A detailed description of the conjoint data is provided in Table 3.

Table 3 Data Description

	Percent off	Limited quantity	Volume discount	Total
Number of respondents	249	233	197	679
Avg. quantity	1.64	1.91	2.14	1.88
Avg. number of variety	0.98	1.09	1.18	1.08
Avg. expense (USD)	4.17	5.28	5.90	5.05
Observation of zeros (%)	87.7	86.4	85.3	86.5
Observation of ones (%)	7.8	8.0	8.1	8.0
Observation of twos (%)	2.7	3.2	4.5	3.4
Observation of threes and above (%)	1.8	2.4	2.1	2.1
Task with no choice (%)	35.8	33.9	28.4	33.0
Task with one variety (%)	42.0	39.1	42.4	41.1
Task with two varieties (%)	14.6	16.8	18.7	16.6
Task with three and above varieties (%)	7.6	10.2	10.5	9.3

Figure 4 (Color online) Example Screenshot of Conjoint Analysis Exercise



3.1. Descriptive Analysis

Our conjoint study asks respondents to indicate their purchases volumetrically. The volumetric nature of our data allows us to compare the raw data to actual market shares. Figure 5 shows a comparison of the observed shares to actual market shares (Beverage Digest 2012) for the top 10 brands. The correlation between the market share and the choice shares is 0.93 indicating that the conjoint survey adequately measures the true relative preferences observed in the marketplace.

Figure 6 displays the distribution of purchase quantities of each brand for the null, unpromoted condition, and each of the three experimental conditions. Each of the promotions leads to increased purchases, although the magnitude of the increase appears to be somewhat small. The apparently small effect of price is due to the large proportion of zeros in the data. That is, most people do not purchase most of the brands. Thus putting a brand on sale that is not preferred appears to have a small effect.

The pattern of purchase volumes also changes with the promotion type. The limited quantity coupon promotion is associated with the greatest increase in volumes for small quantities where the discount is operational. The volume discount promotion is associated with a decrease in the proportion of customers purchasing only one unit, and leads to increases in higher volume purchases. Thus, the raw data show reasonable responses to the conditions.

We begin by fitting a reduced form regression model to our data to illustrate the inadequacy of using a

Figure 5 (Color online) Comparison of Market Shares vs. the Observed Shares in the Data

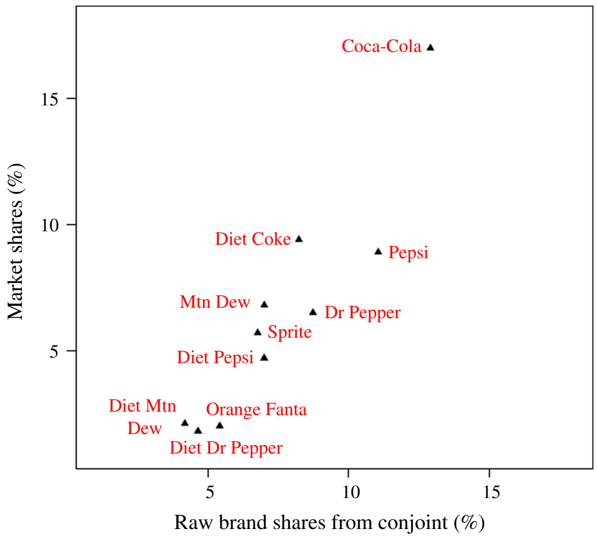
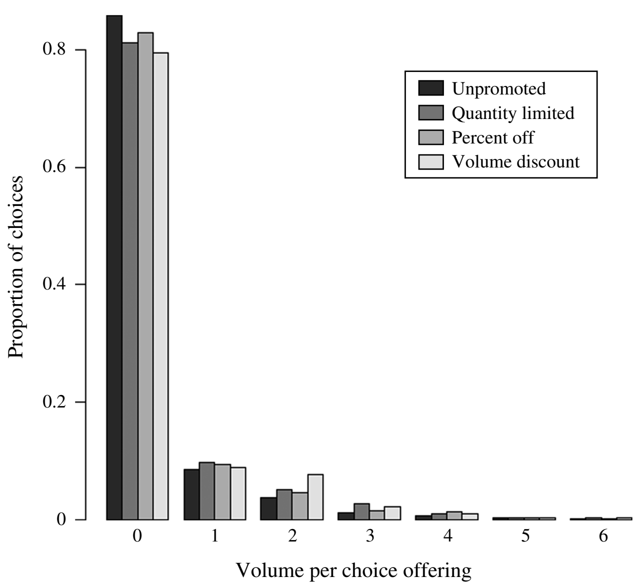


Figure 6 Distribution of Volume Purchase



Note. The figure is truncated to only show volumes fewer than or equal to six. There were few observations with more than six units purchased.

descriptive model. The regression model does not account for competition, the endogeneity of price caused by the nonlinear pricing schemes or the effect of the discount on the budget constraint. Allowing for competition requires a large number of cross-effects among the 48 (16 brands × 3 package sizes) different product offerings.[3] In addition, there is no natural way to represent the nonlinear price schedule in the model without enumerating all possible price combinations among the offerings. We therefore use a simple

[3] There would be 435 cross-effects to estimate to account for competition.

Table 4 Coefficients and Standard Errors from Reduced Form Regression

	Percent off	Limited quantity	Volume discount
	Mean (S.D.)	Mean (S.D.)	Mean (S.D.)
Intercept	0.594 (0.049)	0.650 (0.058)	0.643 (0.058)
Diet Coke	−0.130 (0.048)	−0.250 (0.055)	−0.152 (0.060)
Pepsi	−0.073 (0.044)	−0.059 (0.058)	−0.145 (0.057)
Diet Pepsi	−0.199 (0.046)	−0.206 (0.059)	−0.257 (0.058)
Dr Pepper	−0.153 (0.040)	−0.131 (0.046)	−0.179 (0.051)
Diet Dr Pepper	−0.236 (0.046)	−0.309 (0.055)	−0.337 (0.055)
Mountain Dew	−0.224 (0.043)	−0.227 (0.049)	−0.248 (0.051)
Diet Mountain Dew	−0.311 (0.043)	−0.335 (0.054)	−0.350 (0.054)
7-Up	−0.288 (0.041)	−0.290 (0.050)	−0.296 (0.050)
Diet 7-Up	−0.310 (0.041)	−0.355 (0.050)	−0.339 (0.054)
Sierra Mist	−0.309 (0.041)	−0.356 (0.051)	−0.366 (0.052)
Diet Sierra Mist	−0.314 (0.043)	−0.377 (0.052)	−0.407 (0.052)
Sprite	−0.202 (0.039)	−0.215 (0.048)	−0.278 (0.049)
Sprite Zero	−0.314 (0.041)	−0.367 (0.052)	−0.409 (0.052)
Orange Fanta	−0.280 (0.039)	−0.320 (0.048)	−0.263 (0.049)
Orange Crush	−0.267 (0.040)	−0.307 (0.049)	−0.301 (0.049)
6-pack 24 oz. bottles	0.016 (0.037)	−0.001 (0.036)	0.026 (0.041)
12-pack 12 oz. cans	0.145 (0.041)	0.142 (0.041)	0.114 (0.047)
Price	−0.078 (0.022)	−0.073 (0.024)	−0.059 (0.029)
Promotion dummy	0.110 (0.025)	0.102 (0.026)	0.120 (0.032)
Mean *R*-squared	0.612	0.629	0.622

specification with own-effects only. We fit a Bayesian hierarchical linear model to account for the difference in individual responses. While this model suffers from a number of theoretical disadvantages, it is commonly used in practice for this type of data. The individual model is specified as

$$\text{Volume}_i = \beta_{0i}\text{Brand}_i + \beta_{ps}\text{Package Size}_i + \beta_p\text{Price}_i + \beta_{pr}\text{Promoted}_i + \varepsilon_i, \quad (16)$$

where i indexes the 160 different product offerings from the conjoint survey. The parameter estimates of the model are reported in Table 4.

We compute an R-squared value for each model. Because this is not common practice in Bayesian applications, we describe the process in more detail. There are two complications to computing a measure of in-sample fit in a hierarchical Bayesian paradigm. The first is that with a Bayesian approach we estimate the full distribution of parameter values from the regression model rather than a single point estimate that we get from ordinary least squares (OLS). Rather than display the complete distribution of individual R-squared values we report the mean of the distribution of mean R-squared values for each respondent. The mean R-squared value is 0.612 for the percent off condition, 0.629 for the limited quantity promotion condition, and 0.622 for the volume discount condition.[4]

[4] We also investigated a number of homogenous models. The fit from these models was uniformly poor indicating significant heterogeneity in the individual responses. The R-squared values for the standard OLS model are 0.033, 0.030, and 0.021.

The regression estimates for price and the promotional dummy imply that putting a product on promotion is estimated to have a similar effect as reducing the price of an offering by between \$1.40 and \$2.03. For comparison, the percent off data set offered respondents a 25% flat promotional discount. The average product price was \$3.03 leading to a price discount of \$0.76. This suggests that the promotional elasticities are greater than price elasticities for all three conditions, which is consistent with previous research (Blattberg et al. 1995).

3.2. Empirical Results

Our proposed model of demand described in Equation (15) incorporates the impact of price promotions through the budget set and not through changes to the utility function. We investigate the adequacy of this assumption by including an additional dummy promotional variable when estimating the model for each of the three data sets. That is, instead of g_{li} as described in Equation (12), we add the term $(\beta_i D_i)$, which allows utility to be affected in addition to the budget set. This extended specification allows us to separate the impact of a promotional price change due to the price change from a promotion-induced shift in the utility function (see Inman et al. 1990, Naylor et al. 2006, Wansink et al. 1998 for behavioral work that suggests shifts in utility functions, see Guadagni and Little 1983 for an example of how this is commonly modeled). Including a dummy variable for the promotional effect is a reduced-form method to account for additional effects due to the promotion and which are not captured by the simple change in prices. This could include the effect due to the additional advertising the product on promotion received, any special feature or display that accompanies the promotion or any psychological influences induced by the promotional pricing. To identify the specific mechanism behind the shift in the utility function would require a different parameterization of the utility function or an experiment to rule out alternative explanations and thus is beyond the scope of this paper.

Shifts in the utility function would lead to smooth changes in demand with each respondent experiencing a separate shock when a product is placed on promotion. On the other hand, accounting for quantity-based pricing through the budget set will lead to a discontinuous response in consumers' purchase behavior. It is difficult to determine from the data in our experimental setting whether there are mass build-ups and discontinuities; the additional integer constraint distorts the demand. Because consumers are only allowed to purchase whole number quantities, they must shift purchases from the optimal fractional quantity to the nearest allowable purchase volume. While this prevents one from demonstrating the mass build-ups, the differences in purchase behavior between the promotions conditions can be seen in Figure 6.

Contrary to the simple assumption that a promotion causes a shift in the utility function, each of the different promotion conditions shows a shift in purchase behavior consistent with the economic model of demand. As expected, all of the promotions increased purchase incidence. However the quantity limited promotion led to the greatest increase in small volumes; the volume discount had the greatest impact for larger volumes; and the flat percent off discount had an intermediate impact across all purchase volumes.

Parameter estimates from our proposed model are reported in Table 5. The reported value is the mean of the random-effect specification of our model along with the standard deviation of heterogeneity. The posterior standard deviations of the estimates are approximately equal to 0.15 for the mean parameters. The full covariance matrix of random-effects is available from the authors on request. Overall, we find general consistency of the brand and package size coefficients among the data sets. This indicates that estimates of these aspects of utility are invariant to the promotional type. The promotional dummy variable is, on average, significantly different from zero for only the volume discount data set, and appears to lead to a reduced amount of estimated satiation (γ). That is, the presence of a volume discount promotion leads to an increased demand that is interpreted by the model as a small increase in utility (0.297) and a reduction in the rate of satiation. Expenditure levels (E) are found to be approximately equal across data sets.

We examined a series of alternative models in addition to our proposed model. Table 6 reports model fit statistics (Log Marginal Density (LMD) and R-squared) for (i) models without the promotional dummy variable and (ii) models that assume the flat percent off discount when an irregular quantity-based pricing schedule is placed.[5] It is common in reduced-form models as well as in indirect utility models to apply the percent off discount instead of quantity-based pricing schemes because they do not separate consumers' utility from their constraints. In these models the impact of a temporary price discount is captured through a promotional dummy. This provides a good baseline from which to consider the benefit of accounting for the full price schedule in the budget constraint. We also provide an R-squared measure to easily compare the direct utility models with the reduced-form models presented in §3.1. The measure is calculated in the same manner

[5] This model is similar to the model in Kim et al. (2002) with a log utility function rather than a power utility function to ensure model consistency.

Table 5 **Posterior Mean and Standard Deviation of Heterogeneity of Random-Effects**

	Percent off		Limited quantity		Volume discount	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Intercept	−4.463	(0.192)	−4.158	(0.181)	−3.909	(0.156)
Diet Coke	−0.987	(0.155)	−1.066	(0.170)	−0.892	(0.176)
Pepsi	−0.260	(0.138)	−0.199	(0.137)	−0.620	(0.149)
Diet Pepsi	−1.045	(0.165)	−0.942	(0.187)	−1.290	(0.205)
Dr Pepper	−0.976	(0.134)	−0.606	(0.115)	−0.771	(0.136)
Diet Dr Pepper	−1.517	(0.181)	−1.300	(0.174)	−1.390	(0.166)
Mountain Dew	−1.664	(0.163)	−0.972	(0.155)	−1.069	(0.137)
Diet Mountain Dew	−1.711	(0.142)	−1.590	(0.168)	−1.516	(0.162)
7-Up	−1.201	(0.118)	−1.009	(0.127)	−0.908	(0.139)
Diet 7-Up	−1.539	(0.161)	−1.271	(0.143)	−1.410	(0.166)
Sierra Mist	−1.642	(0.150)	−2.322	(0.166)	−1.815	(0.190)
Diet Sierra Mist	−1.767	(0.180)	−2.100	(0.186)	−2.659	(0.220)
Sprite	−0.699	(0.110)	−0.558	(0.114)	−0.740	(0.118)
Sprite Zero	−1.516	(0.143)	−1.432	(0.143)	−1.728	(0.147)
Orange Fanta	−1.389	(0.115)	−1.617	(0.143)	−1.515	(0.189)
Orange Crush	−1.539	(0.131)	−1.159	(0.128)	−1.714	(0.175)
6-pack 24 oz. bottles	−0.176	(0.136)	0.054	(0.112)	0.261	(0.128)
12-pack 12 oz. cans	1.161	(0.124)	1.287	(0.126)	1.231	(0.137)
Promotion dummy	0.060	(0.057)	0.045	(0.058)	0.297	(0.063)
$\log(\gamma)$	2.085	(0.195)	1.603	(0.234)	1.430	(0.196)
$\log(E)$	3.368	(0.082)	3.420	(0.084)	3.341	(0.088)
LMD (log marginal density)	−10,660.7		−10,875.0		−10,677.8	
Mean R-squared	0.715		0.722		0.704	

Table 6 **Comparison of Model With and Without Quantity Based Pricing**

		Promotional effect	LMD	R-squared
Percent off	w/o quantity pricing	0.060 (0.057)	−10,660.7	0.715
	w/o promotional dummy	0 (n/a)	−10,836.0	0.698
	w/both (as proposed)	0.060 (0.057)	−10,660.7	0.715
Limited quantity	w/o quantity pricing	0.207 (0.058)	−11,007.9	0.719
	w/o promotional dummy	0 (n/a)	−11,095.4	0.702
	w/both (as proposed)	0.045 (0.058)	−10,875.0	0.722
Volume discount	w/o quantity pricing	0.319 (0.060)	−10,760.3	0.700
	w/o promotional dummy	0 (n/a)	−10,981.5	0.676
	w/both (as proposed)	0.297 (0.063)	−10,677.8	0.704

and captures the correlation between the fitted values from the model and the observed demand.

We find that the absence of a promotional dummy variable in the model specification leads to a reduced fit in the limited quantity and volume discount case although the coefficient for the promotional dummy is only significant in the volume discount case. This suggests that there may be substantial heterogeneity in consumer responses to promotional activities with some consumers responding to the promotion to a greater degree than can be expected from just the economic price change. Additional research is warranted to determine why different promotional offers influence different consumers to such a high degree.

Table 6 also compares the proposed model to the horizontal variety model of Kim et al. (2002) identified as the model without quantity pricing. The fit of the model without the quantity-based pricing is worse in the limited quantity and volume discount case. (The model is mathematically identical in the flat percent off promotion case.) In addition, the promotional dummy coefficients are overestimated when quantity-based pricing is not accounted for, suggesting that the impact of the price promotion through the budget constraint is misattributed to the direct impact of the price promotion on the utility. This would lead to systematic bias in counterfactual analysis.

All of the direct utility models show an improvement in in-sample fit over the reduced-form models. The direct utility models improve the R-squared values by about 15% from approximately 0.62 to 0.71 with a minimal change in the total number of parameters estimated.

Table 7 Optimal Promotion Types

	Promotional effect = 0		Promotional effect from posterior	
Optimal promotion type	No. of cust.	% of cust.	No. of cust.	% of cust.
No discount	168	24.7	47	6.9
15% off Coke products	117	17.2	54	8.0
25% off Coke products	7	1.0	3	0.4
35% off Coke products	0	0.0	0	0.0
15% off Coke products (limit 3)	100	14.7	55	8.1
25% off Coke products (limit 3)	7	1.0	4	0.6
35% off Coke products (limit 3)	2	0.3	1	0.1
15% off Coke products (after 3)	215	31.7	422	62.2
25% off Coke products (after 3)	51	7.5	81	11.9
35% off Coke products (after 3)	12	1.8	12	1.8
Total	679	100.0	679	100.0

4. Counterfactual Analysis

We demonstrate one potential application of the model based on individually customized marketing messages to consumers. As data becomes more accessible, companies target consumers in increasingly smaller segments. In the limit, companies can target individual consumers. This level of customization is already present in Internet marketing via a process called retargeting where customers are shown advertisements for products that they have recently viewed on a retail website (Lambrecht and Tucker 2013). It is also prevalent in database marketing where a combination of increased data availability and inexpensive customized printing have enabled stores to send out customized sales flyers (see Duhigg 2012 for an example).

We show how this targeting can be expanded to include customized pricing offers. The analysis presented in this section is similar to an analysis presented in Rossi et al. (1996). They demonstrate a model that can be used to determine the discount level that should be offered to consumers. Their analysis, however, simply investigated the depth of the discount offered and did not consider the possibility of offering different pricing schedules to each consumer. Because different discounts can have a different promotional impact, discount types can be more important than discount levels.

For demonstration purposes, we optimize the pricing of Coca-Cola products in the study. There are five Coca-Cola brands in the study: Coca-Cola Classic, Diet Coke, Sprite, Sprite Zero, and Orange Fanta. We consider nine different promotion type and level combinations together with a base condition where no promotion is offered. For the base condition we assume that each firm prices all brands in their portfolio the same for a given size. For simplicity, we assume that Coca-Cola products are offered at \$3.99, \$3.49, and \$1.59 for the cans, bottles, and 2-liter bottle, respectively. PepsiCo products are 10 cents below the prices of Coca-Cola and Dr Pepper/Snapple Group products are 15 cents below Coca-Cola prices. Because we do not have firm specific costs, we assume that the firms enjoy a fixed margin of 50% of the base retail price. Different margins lead to qualitatively similar results. The specific levels tested are presented in Table 7.

The results from §3 form the basis of the counterfactual analysis with one minor adjustment. The volume discount cell had a positive promotional effect on utility while the other two cells did not. Because of the design of the experiment, each respondent only saw one of the three promotion types. If the dummy variable representing the promotion condition differs for each type of promotion, it is necessary to collect data on all of the promotions for each respondent to estimate a promotion type specific dummy. The experimental setting we designed does not have this individual level variation in promotional type and as such we run two scenarios that assume a given promotional response. The first scenario assumes no promotional response beyond the change in pricing captured in the budget constraint (e.g., we set the promotional dummy to be 0 for all promotion conditions). This represents a conservative case since the promotional dummy is expected to be positive. The second case uses the mean promotional dummy of each condition as a respondent's promotional dummy. This represents a more realistic case as it accounts for an additional promotional response beyond the change in pricing.

We predict the purchases and profitability for each consumer under each of the 10 scenarios to determine which promotion to offer to each customer. Customers are offered the promotion that will lead to the greatest profitability for Coca-Cola. The distribution of individual promotions is presented in Table 7. Note that even in the conservative case the most common offer is the volume discount promotion with a low discount level.

Table 8 presents the rise in profit that could be expected over the no promotion scenario. The targeting strategy presents a 20.41% rise over the no promotion strategy in the conservative case and a 34.21% rise when including the additional promotional effects.

Table 8 Profit for Various Promotional Strategies

	Promotional effect = 0		Promotional effect from posterior	
	Profit ($)	Lift (%)	Profit ($)	Lift (%)
Base case (no promotion)	2,027.59	—	2,026.91	—
Best uniform promotion (15% volume discount)	2,186.73	7.85	2,584.28	27.50
Full customization	2,441.45	20.41	2,720.32	34.21

If targeted promotions are not possible, volume discounts still represent a more profitable strategy than a constant promotion. The 15% volume discount would lead to a 7.85% rise in profit under the conservative case and a 27.50% rise when we allow for the positive shift in utility due to the price promotion.[6]

5. Conclusion

We present a general method for modeling demand in the presence of nonlinear pricing and quantity discounts. Quantity-based pricing is a common pricing technique that has not been addressed well in the empirical demand literature. Reduced form and indirect utility models are not well suited to modeling quantity-based price discounts. The presence of price discounts can create kinks and discontinuities in the budget set, leading to a build-up of likelihood mass in the interior of the demand space. We show how to address this complexity using a direct utility model without relying on FOCs. The proposed method is general in that it can be applied to a variety of pricing schedules and utility functions.

Quantity-based pricing as discussed in this paper is different from the traditional retail format of pricing different package sizes at different per unit prices (Allenby et al. 2004). When combining different package sizes and formats, it is not clear that the differential demand for the package sizes is due solely to the changes in volumes and unit prices. This paper treats each package (i.e., 12-pack of cans) as a single unit and applies the quantity discount across multiple package purchases. This avoids the problem of determining what the appropriate unit is within a package as well as the difference in units (i.e., bottles versus cans) that occur across package sizes.

We apply this model to an online conjoint experiment of soda preferences. Volumetric demand data is collected for individual respondents under no-promotion and promotion conditions. This format was used to verify that capturing the quantity-based price change is adequate to explain purchasing behavior. The evidence suggests that consumers not only shift their behavior according to the price schedule but also may experience a shift in their utility function when faced with a promotion. There is significant heterogeneity in the degree to which respondents shift their utility function. Additional research is necessary to fully understand the relationship between these two types of behavioral responses and the way that different consumers respond to these price promotions. The method presented in this paper is a foundation for this additional work.

The direct utility model also allows for individual estimates of the impact of promotional activities. We demonstrate the value of these individual estimates in our counterfactual analysis. The profit maximizing promotional strategy is heterogeneous, with some respondents receiving no promotion and others receiving a large price discount. This is true even without taking into account the different promotional sensitivities that consumers may exhibit.

We recognize that it is not currently common practice for firms to offer this level of price customization. The full optimization approach requires knowledge of customer-specific utility functions and parameters. Firms can track consumers as they are shopping and deliver customized promotions (see Weber 2004 for an example). Our model could easily be applied to existing purchase databases that stores maintain through the use of store loyalty cards and are delivered to customers through customized alerts on their mobile phones.

Even if firms are unable to engage in consumer-specific targeting, they can still benefit from the methodology presented in this paper. Our methodology can be used to develop traditional and new pricing strategies and design promotions targeted at specific behavioral groups of consumers such as switchers or those with low price sensitivity.

Although the model is useful as presented, there are a number of ways that future research could extend the model and make it more flexible. The model does not currently account for temporal dynamics in the purchasing process. In the face of promotions, consumers may purchase storable products in anticipation of future price increases. This would lead them to decrease future purchases due to the stock of products they have on hand. This does not involve conjoint style data such as that used in this paper since the task is designed to minimize the presence of strategic buying, however this could significantly influence the results obtained from actual purchase data. The addition of temporal substitution into the model is nontrivial due to potential changes in the budget constraint between time periods. While it is reasonable to assume that the respondent knows the current budget, the future budget is only known with expectation.

[6] The profit presented does not attempt to correct for market size and is thus based on the 679 respondents in the study.

Another limitation of the model is the absence of direct substitution effects. In the model, substitution between goods happens only through the income effect due to the nonlinear outside good. This means that all of the inside goods in the model have a proportional substitution pattern. Additional research is necessary to extend the model to handle direct substitution between the goods. The primary challenge to incorporating a model with flexible substitution patterns is the number of parameters needed to estimate unique cross-product substitution effects. This dramatically expands the number of parameters to be estimated especially for attribute-based models such as the one used in the paper. For example, there are 48 unique products represented in the experiment presented in the paper. This leads to over 1,000 additional parameters that would need to be estimated to capture the substitution effects. It is not possible to identify the substitution patterns with any degree of accuracy with the limited data series available. Additional modeling work or restructured utility functions are necessary to capture these substitution patterns.

Despite these limitations, our model provides a reasonable representation of the consumers purchasing process in the face of quantity discounts. These types of discounts are commonly seen in the purchasing process and require special modeling considerations. Direct utility methods provide a natural way for researchers to incorporate these nonlinear price schedules into consumer demand models.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/mksc.2015.0948.

Acknowledgments

The authors thank the associate editor and two anonymous reviewers for their helpful comments.

Appendix A. Discrete Demand Estimation

A consumer's utility function can be written in terms of the inside goods by substituting the binding budget constraint[7] for the outside good

$$\begin{aligned} U^*(x^*_{1t}, \ldots, x^*_{It}) &= \sum_{i=1}^{I} \psi_{it} \log(\gamma x^*_{it} + 1) \\ &\quad + \log\Big(E - \sum_{i=1}^{I}[p_{lit} \min(x^*_{it}, \tau_i) + p_{hit} \max(0, x^*_{it} - \tau_i)]\Big) \\ &\text{where } \psi_{it} = \exp(a_i\beta + \varepsilon_{it}) \quad (\text{i.e., } \psi^*_{it} = a_i\beta + \varepsilon_{it}) \\ &\qquad \varepsilon_{it} \sim \mathrm{N}(0, 1). \end{aligned}$$

[7] A budget constraint is always binding when optimal decisions are made.

Because solving a nonlinear integer programming problem is known to be NP-hard (Wolsey 1998), the comparisons with all of the feasible points are required. The sufficient conditions for optimality are given as follows:

$$U^*(x^*_{1t}, \ldots, x^*_{It}) \geq \max\Big\{U^*(x_{1t}, \ldots, x_{It}) \Big| E - \sum_{i=1}^{I}[p_{lit}\min(x_{it}, \tau_i) + p_{hit}\max(0, x_{it} - \tau_i)] \geq 0, x_{it} \in \{0, 1, \ldots\}\Big\}.$$

Note that the number of feasible points increases exponentially as the number of goods increases. Lee and Allenby (2014) propose to compare the observed optimal decision only with the adjacent feasible points, instead of the full comparisons.

$$U^*(x^*_{1t}, \ldots, x^*_{It}) \geq \max\{U^*(x^*_{1t} + \Delta_1, \ldots, x^*_{It} + \Delta_I) \mid (x^*_{1t} + \Delta_1, \ldots, x^*_{It} + \Delta_I) \in F\}_{\Delta_i \in \{-1,0,1\}}.$$

Because the comparisons with the adjacent points are not sufficient but necessary conditions for optimality, we can conduct the robustness check of the estimation results by increasing the scope of the adjacent points (e.g., $\Delta_i \in \{-2, -1, 0, 1, 2\}$). When the presence of a quantity-based price promotion leads to an additional irregularity due to the kinks and points of discontinuity in the budget constraint, we can include them as additional comparison points for optimality.

We want to infer about all of the unknowns: β, γ^*, E^*, and introduce $\{\psi^*_{1t}, \ldots, \psi^*_{It}\}_{t=1,\ldots,T}$ as augmented variables (γ^*, E^*, ψ^*_{it} denote $\log(\gamma)$, $\log(E)$, $\log(\psi_{it})$, respectively). The joint posterior distribution of the parameters and augmented variables is proportional to the likelihood times the prior

$$\begin{aligned} &p(\beta, \gamma^*, \{\psi^*_{1t}, \ldots, \psi^*_{It}\}_{t=1,\ldots,T} \mid \{x_{1t}, \ldots, x_{It}\}_{t=1,\ldots,T}) \\ &\quad \propto \prod_{t=1}^{T} l(x_{1t}, \ldots, x_{It} \mid \psi^*_{1t}, \ldots, \psi^*_{It}, \gamma^*, E^*) \\ &\quad \times \prod_{i=1}^{I}\prod_{t=1}^{T} p(\psi^*_{it} \mid \beta) \times p(\beta) \times p(\gamma^*) \times p(E^*). \end{aligned}$$

Given $\{\psi^*_{1t}, \ldots, \psi^*_{It}, \gamma^*, E^*\}$, the likelihood of $(x_{1t}, \ldots, x_{It})$ is an indicator function whose value equals one when all of the optimality conditions are satisfied and zero otherwise. In addition, β is conditionally independent of the data given $\{\psi^*_{1t}, \ldots, \psi^*_{It}\}_{t=1,\ldots,T}$. Therefore, the full conditional distributions can be expressed as

$$\begin{aligned} &p(\psi^*_{it} \mid else) \propto l(x_{1t}, \ldots, x_{It} \mid \psi^*_{1t}, \ldots, \psi^*_{It}, \gamma^*, E^*) \times p(\psi^*_{it} \mid \beta), \\ &p(\beta \mid else) \propto \prod_{i=1}^{I}\prod_{t=1}^{T} p(\psi^*_{it} \mid \beta) \times p(\beta), \\ &p(\gamma^* \mid else) \propto \prod_{t=1}^{T} l(x_{1t}, \ldots, x_{It} \mid \psi^*_{1t}, \ldots, \psi^*_{It}, \gamma^*, E^*) \times p(\gamma^*), \\ &p(E^* \mid else) \propto \prod_{t=1}^{T} l(x_{1t}, \ldots, x_{It} \mid \psi^*_{1t}, \ldots, \psi^*_{It}, \gamma^*, E^*) \times p(E^*). \end{aligned}$$

The conditional distribution of ψ^*_{it} is a truncated normal distribution whose upper and lower truncation points are determined by the comparisons with adjacent points given $\{\psi^*_{jt}\}_{j\neq i}$. Given $\{\psi^*_{1t}, \ldots, \psi^*_{It}\}_{t=1,\ldots,T}$, β can be generated from a simple Bayesian regression model. However, we cannot draw directly from the posterior distributions for γ^*, E^*. Therefore, we use the Metropolis-Hastings algorithm for drawing γ^*, E^*. Consequently, we can simulate all of the unknowns from the joint posterior distribution by using a Gibbs-type sampler and following a sequence of steps described below:

Step 1. Set initial values for β, γ^*, E^*

Step 2. Set initial values for $\{\psi_{1t}^*, \ldots, \psi_{It}^*\}_{t=1,\ldots,T}$. By solving the Kuhn-Tucker conditions with given β, γ^*, E^*.

Step 3. Draw $\{\psi_{1t}^*, \ldots, \psi_{It}^*\}_{t=1,\ldots,T}$

$$\psi_{it}^* \mid \{\psi_{jt}^*\}_{j\neq i}, \beta \sim \mathrm{N}(a_i\beta, 1) \times I(lb_{it} \leq \psi_{it}^* \leq ub_{it})$$

where lb_{it}, ub_{it} denote the lower and upper truncation points for ψ_{it}^*. The function $I(\cdot)$ is an indicator function with a value of 1 if ψ_{it}^* is between lb_{it} and ub_{it} and 0 otherwise.

Step 4. Draw β

$$\beta \mid \{\psi_{1t}^*, \ldots, \psi_{It}^*\}_{t=1,\ldots,T}, \beta_0, \sigma_0$$
$$\sim \mathrm{N}\left(\left(\frac{\beta_0}{\sigma_0^2} + \sum_{i=1}^{I}\sum_{t=1}^{T}\psi_{it}^*\right)\cdot\left(\frac{1}{\sigma_0^2}+IT\right)^{-1}, \left(\frac{1}{\sigma_0^2}+IT\right)^{-1}\right).$$

The prior distribution for β is $N(\beta_0, \sigma_0^2)$.

Step 5. Draw γ^* using a Metropolis-Hastings algorithm with random-walk chain.

Step 6. Draw E^* using a Metropolis-Hastings algorithm with random-walk chain.

Step 7. Repeat Step 3 through Step 6 at each iteration of the MCMC. The draws at the previous iteration become the new initial values.

Appendix B. Graphical Appendix

B.1. Example Conjoint Tasks

Figure B.1 **(Color online) Example of No Discount Question**

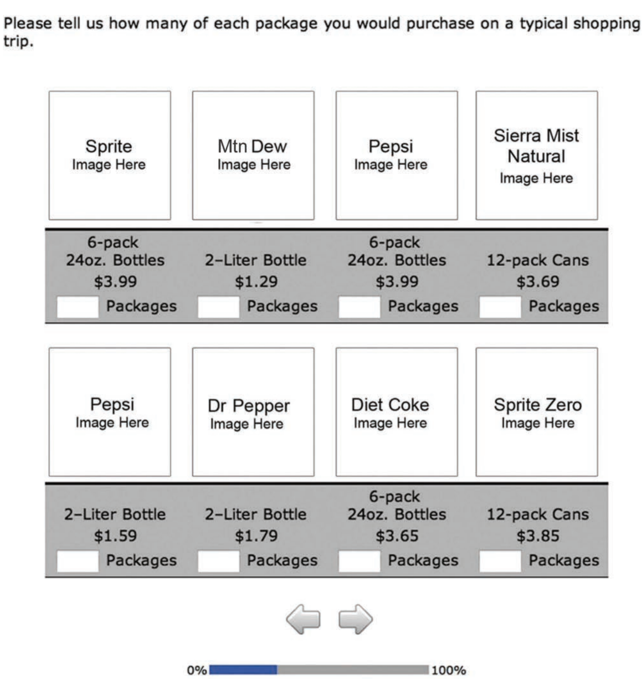


Figure B.2 **(Color online) Example of Limited Quantity Discount Question**

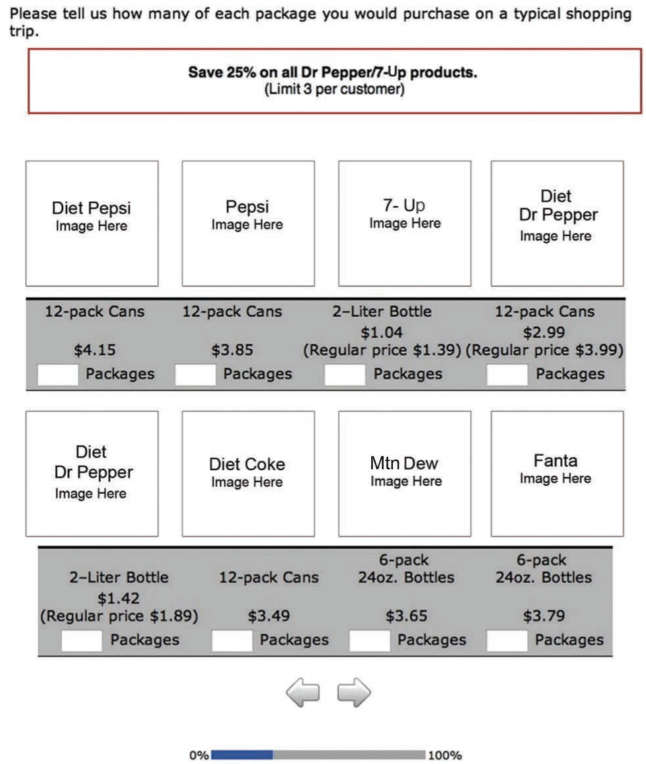
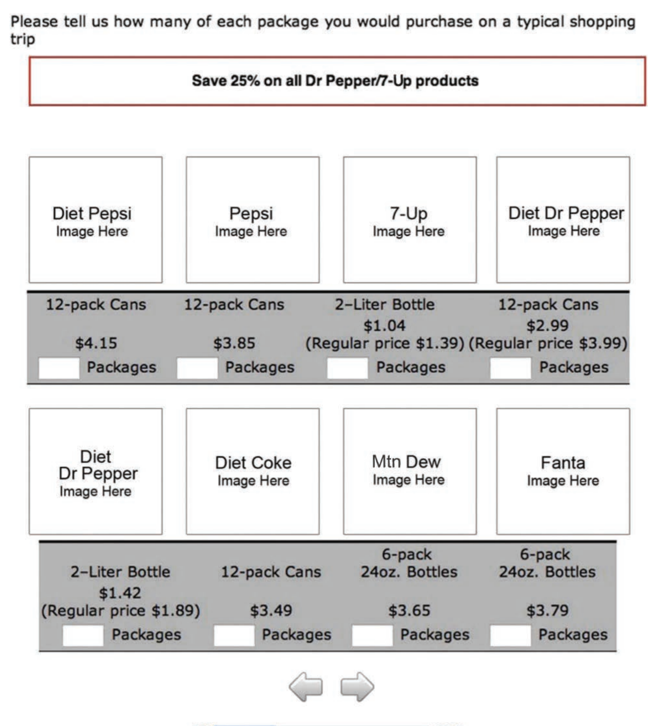


Figure B.3 **(Color online) Example of Percent Off Discount Question**



Appendix C. Covariance Matrices

Table C.1 Posterior Means of Diagonal Elements of Variance-Covariance Matrix

	Percent off	Limited quantity	Volume discount
Intercept	7.978	7.469	4.727
Diet Coke	5.441	5.713	4.776
Pepsi	3.499	3.786	3.512
Diet Pepsi	6.145	6.217	5.849
Dr Pepper	2.929	2.808	2.436
Diet Dr Pepper	6.507	6.735	4.424
Mountain Dew	5.447	3.574	2.915
Diet Mountain Dew	5.004	5.534	4.512
7-Up	2.507	2.659	2.432
Diet 7-Up	3.747	3.616	4.335
Sierra Mist	4.519	5.174	4.474
Diet Sierra Mist	5.955	5.743	6.629
Sprite	2.145	2.407	2.226
Sprite Zero	3.940	3.947	3.562
Orange Fanta	2.480	3.640	3.665
Orange Crush	3.994	3.330	4.290
6-pack 24 oz. bottles	3.219	2.535	2.601
12-pack 12 oz. cans	3.533	3.405	3.381
Promotion dummy	0.509	0.523	0.511
$\log(\gamma)$	8.440	11.475	6.616
$\log(E)$	1.461	1.387	1.469

References

Allenby GM, Shively TS, Yang S, Garratt MJ (2004) A choice model for packaged goods: Dealing with discrete quantities and quantity discounts. *Marketing Sci.* 23(1):95–108.

Beverage Digest (2012) U.S. beverage results for 2011. *Beverage Digest* 61(6):1–2. http://www.beverage-digest.com/pdf/top-10_2012.pdf.

Bhat CR (2005) A multiple discrete-continuous extreme value model: Formulation and application to discretionary time-use decisions. *Transportation Res. Part B: Methodological* 39(8):679–707.

Blattberg RC, Briesch RA, Fox EJ (1995) How promotions work. *Marketing Sci.* 14(3, Suppl):G122–G132.

Deaton A, Muellbauer J (1980) *Economics and Consumer Behavior* (Cambridge University Press, New York), 15–16.

Duhigg C (2012) How companies learn your secrets. *New York Times* (February 16). http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html.

Guadagni PM, Little JDC (1983) A logit model of brand choice calibrated on scanner data. *Marketing Sci.* 2(3):203–238.

Howell JR, Allenby GM (2014) Choice models with fixed costs. Working paper 1-40, Pennsylvania State University, University Park.

Inman JJ, McAlister L, Hoyer WD (1990) Promotion signal: Proxy for a price cut? *J. Consumer Res.* 17(1):74–81.

Iyengar R, Jedidi K (2012) A conjoint model of quantity discounts. *Marketing Sci.* 31(2):334–350.

Kim J, Allenby GM, Rossi PE (2002) Modeling consumer demand for variety. *Marketing Sci.* 21(3):229–250.

Lambrecht A, Tucker C (2013) When does retargeting work? Information specificity in online advertising. *J. Marketing Res.* 50(5): 561–576.

Lee S, Allenby GM (2014) Modeling indivisible demand. *Marketing Sci.* 33(3):364–381.

Lee S, Kim J, Allenby GM (2013) A direct utility model for asymmetric complements. *Marketing Sci.* 32(3):454–470.

Mas-Colell A, Whinston MD, Green JR (1995) *Microeconomic Theory* (Oxford University Press, New York), 956–964.

McFadden D (1986) The choice theory approach to market research. *Marketing Sci.* 5(4):275–297.

Mehta N, Chen X, Narasimhan O (2010) Examining demand elasticities in Hanemann's framework: A theoretical and empirical analysis. *Marketing Sci.* 29(3):422–437.

Naylor RW, Raghunathan R, Ramanathan S (2006) Promotions spontaneously induce a positive evaluative response. *J. Consumer Psychol.* 16(3):295–305.

Rossi PE, McCulloch RE, Allenby GM (1996) The value of purchase history data in target marketing. *Marketing Sci.* 15(4):321–340.

Satomura T, Brazell JD, Allenby GM (2011a) Choice models for budgeted demand and constrained allocation. Working paper 1-31, Ohio State University, Columbus.

Satomura T, Kim J, Allenby GM (2011b) Multiple-constraint choice models with corner and interior solutions. *Marketing Sci.* 30(3): 481–490.

Villas-Boas JM, Winer RS (1999) Endogeneity in brand choice models. *Management Sci.* 45(10):1324–1338.

Wales TJ, Woodland AD (1983) Estimation of consumer demand systems with binding non-negativity constraints. *J. Econometrics* 21(3):263–285.

Wansink B, Kent RJ, Hoch SJ (1998) An anchoring and adjustment model of purchase quantity decisions. *J. Marketing Res.* 35(1): 71–81.

Weber T (2004) The future of shopping. *BBC News* (May 18). http://news.bbc.co.uk/2/hi/business/3712261.stm.

Wolsey LA (1998) *Integer Programming*, Vol. 52, Wiley Series Discrete Math. Optim. (John Wiley & Sons, New York), 81–90.