MARKETING SCIENCE
Vol. 23, No. 1, Winter 2004, pp. 95–108
ISSN 0732-2399 | EISSN 1526-548X | 04 | 2301 | 0095
DOI 10.1287/mksc.1030.0022
© 2004 INFORMS

A Choice Model for Packaged Goods: Dealing with Discrete Quantities and Quantity Discounts

Greg M. Allenby
Fisher College of Business, Ohio State University, 2100 Neil Avenue, Columbus, Ohio 43210, allenby.1@osu.edu
Thomas S. Shively
Department of Management Science and Information Systems, The University of Texas at Austin, Austin, Texas 78712, shively@mail.utexas.edu
Sha Yang
Stern School of Business, 44 W. 44th Street, New York, New York 10012, shayang@stern.nyu.edu
Mark J. Garratt
Strategic Modeling and Forecasting, Miller Brewing Company, 3939 W. Highland Boulevard, Milwaukee, Wisconsin 53201, mgarratt@mbco.com

Utility maximizing solutions to economic models of choice for goods with either discrete quantities or nonlinear prices cannot always be obtained using standard first-order conditions such as Kuhn-Tucker and Roy's identity. When quantities are discrete, there is no guarantee that derivatives of the utility function are equal to derivatives of the budget constraint. Moreover, when prices are nonlinear, as in the case of quantity discounts, first-order conditions can be associated with the minimum rather than the maximum value of utility. In these cases, the utility function must be directly evaluated to determine its maximum. This evaluation can be computationally challenging when there exist many offerings and when stochastic elements are introduced into the utility function. In this paper, we provide an economic model of demand for substitute brands that is flexible, parsimonious, and easy to implement. The methodology is demonstrated with a scanner panel data set of light-beer purchases. The model is used to explore the effects of price promotions on primary and secondary demand, and the utility of product assortment.

Key words: Bayesian analysis; econometric models; pricing research; product management
History: This paper was received June 25, 2001, and was with the authors 3 months for 3 revisions; processed by Pradeep Chintagunta.

1. Introduction

Economic models of brand choice must deal with two complexities when applied to the study of packaged goods. First, consumers are restricted to purchasing discrete quantities of brands. Demand for brands is defined on a grid of available brand-pack combination and not on the entire set of real numbers. Second, the unit price of a brand often depends on the quantity purchased. Quantity discounts are typically available where the per-unit price declines with the size of the package. These two complexities invalidate the use of first-order conditions to identify utility maximizing solutions. First-order conditions may not hold at the available package sizes. Furthermore, constraints imposed by the nonlinear budgetary allotment can lead to first-order conditions that identify the minimum rather than the maximum utility. In this paper, we propose a random utility choice model capable of dealing with discrete quantities and quantity discounts.

Quantity discounts are common in marketing. They occur with packaged goods when the unit price declines with larger quantities, and with services when the price declines as usage increases (e.g., cell phones). Quantity discounts afford manufacturers the ability to price discriminate between high-volume and low-volume users (see Dolan 1987). The depth of price promotions for a particular quantity results in temporary substitution away from other brands and an increase in demand for the product category. An important issue in managing discounts is in determining the expected increase in sales from increased product usage versus substitution from competitive brands. This requires a comprehensive model that can handle a large number of distinct brand-pack combinations that are typically available in many product categories.

Researchers in marketing who are studying consumer demand for packaged goods have tended to ignore the complexities associated with quantity discounts, and have used one of three strategies for deal-

Table 1 Approaches to Dealing with Nonlinear Prices

Treat Different Brand/Size Combinations as Different Choice Alternatives
Guadagni and Little (1983)
Chintagunta (1992)
Fader and Hardie (1996)
Montgomery (1997)
Kalyanam and Putler (1997)
Bell and Lattin (1998)
Chintagunta (1998)
Silva-Risso et al. (1999)
Chang et al. (1999)
Andrews and Manrai (1999)
Focus Analysis on One Package Size per Brand
Allenby (1989)
Blattberg and Wisniewski (1989)
Allenby and Rossi (1991)
Bucklin and Lattin (1991)
Fader et al. (1992)
Chintagunta (1993)
Allenby and Lenk (1994)
Rossi et al. (1996)
Kamakura et al. (1996)
Ainslie and Rossi (1998)
Allenby and Rossi (1999)
Kim et al. (2002)
Employ Average Per-Unit Pricing
Neslin et al. (1985)
Bolton (1989)
Chiang (1991)
Pedrick and Zufryden (1991)
Jain and Vilcassim (1991)
Chiang (1995)
Erdem (1996)
Papatla (1996)
Dillon and Gupta (1996)
Manchanda et al. (1999)
Kadiyali et al. (2000)

ing with unit prices that depend on the purchase quantity. Table 1 provides a selective review of the literature. The first approach to dealing with nonlinear prices is to explode the number of choice alternatives and model each brand-size combination as a distinct choice alternative. In this approach, consumers are assumed to derive utility from having the brand bundled into different quantities. Although various parameters and constraints could, in theory, be introduced into the analysis so that the estimated parameters exhibit desirable properties (e.g., diminishing returns to quantity), a drawback is that it conditions on the product class expenditure and therefore does not provide insight into the trade-off between the product class and other goods. This approach confounds the expenditure decision for the product category and utility for the brand. A consumer who regularly purchases a particular quantity will be seen as having low utility for larger quantities. However, this may be because of the consumer's budget constraint and not because of a lack of preference for larger quantities. The second approach to dealing with nonlinear prices has been to restrict analysis to a particular package size. Whereas such analysis provides valid measures of brand preference, it usually rules out the study of purchase quantity and category expenditure. The third approach uses an average per-unit pricing variable in the analysis. The validity of this approach depends on the degree to which the actual price schedule is linear, which is often not the case.

In this paper, we propose a random utility model for consumer choice that deals with discrete quantities and quantity discounts. The model does not treat different brand-size combinations as different multinomial choice alternatives. Instead, we derive the likelihood specification for the observed demand data from more fundamental assumptions about random utility. This approach results in a parsimonious likelihood specification that facilitates the study of consumer demand across package sizes and does not assume that prices are linearly related to quantity. Our model is not affected by the presence of quantity discounts that are frequently encountered with packaged goods data and deals directly with constraints imposed by discrete package sizes.

In §2, we discuss problems encountered with discrete quantities and quantity discounts encountered in the study of packaged goods. Our approach to dealing with these issues in the context of a brand-choice model is discussed in §3. We note that our model can accommodate the presence of quantity discounts, discrete quantities, or both, and therefore has wide application. Furthermore, both standard linear utility and nonhomothetic utility structures (Allenby and Rossi 1991) can be incorporated into the framework. In §4, we describe a scanner panel data set of light-beer purchases in which there are more than 80 brand-pack combinations available to consumers in the market. Parameter estimates of our parsimonious demand model are discussed in §5. Pricing implications are explored in §6, and conclusions are offered in §7.

2. Problems Encountered with Discrete Quantities and Quantity Discounts

Consider a brand-choice model arising from a linear utility structure ($u(x) = \psi' x$), where the price of the brand depends on quantity. In this case, the vector of marginal utility, ψ, is constant, and utility, u, is maximized subject to the budget constraint $\sum_{k=1}^{K} p_k(x_k) < y$, where $p_k(x_k)$ is the price of x_k units of brand k. Assume that the price function $p_k(x_k)$ reflects quantity discounts. For example, the price schedule could be a marginally decreasing function of quantity:

$$p(x_k) = x_k^{1/a}; \quad a > 1; \ x \geq 0. \tag{1}$$

The identification of the utility maximizing solution using first-order conditions typically proceeds by differentiating the auxiliary function:

$$L = \psi' x - \lambda\left(\sum_{k=1}^{k} p_k(x_k) - y\right), \qquad (2)$$

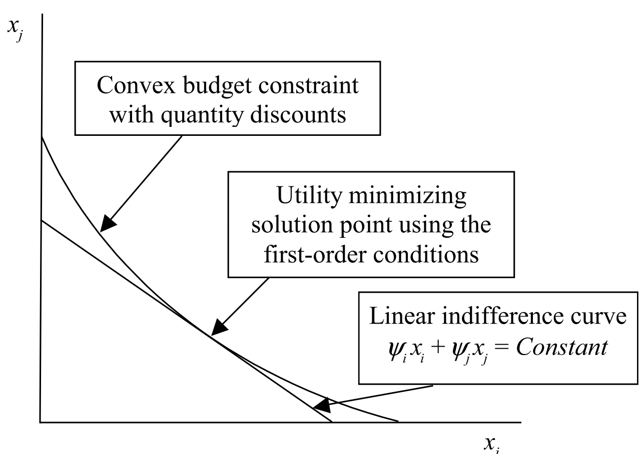
setting first derivatives to zero, and solving for the values of x_k, at which the vector of marginal utility is tangent to the budget curve. We note immediately that, if demand is restricted to available pack sizes, first-order conditions will not necessarily apply at the observed quantity demanded. However, even when quantity is not restricted to a fixed number of values, the use of first-order conditions can lead to a solution identifying a point of utility minimization, not utility maximization. As illustrated in Figure 1 for the case of linear utility and price quantity discounts, the point of tangency is not associated with the utility maximizing solution. Utility is greater at the intersection of the budget curve and either axis.

The reason that first-order conditions are associated with a point of utility minimization rather than utility maximization is because the auxiliary function in Equation (2) is convex, or has a positive second derivative. The second derivative of Equation (2), when prices follow Equation (1), can be shown to be equal to:

$$\frac{\partial^2 L}{\partial x_i^2} = -\lambda\left(\frac{1}{a}\right)\left(\frac{1-a}{a}\right)x_i^{(1-2a)/a}, \qquad (3)$$

which is positive for $a > 1$ and $x > 0$. More generally, whereas the concavity of the auxiliary function is guaranteed when the utility function is concave and the budget constraint is linear, this property is not necessarily present when prices are nonlinear. It is certainly not true when the utility function is linear and the budget constraint is convex. The concavity of the auxiliary function is needed for many identities in economics to be associated with utility maximization, including the Envelope Theorem and Roy's Identity (see Sydsaeter et al. 1999).

Figure 1 Illustration of First-Order Conditions and Utility Minimization



In the next section, we outline an approach to modeling consumer brand choice for packaged goods that do not rely on first-order conditions. Our approach involves the direct evaluation of the utility function. We show that this evaluation can be efficiently carried out for near-perfect substitute goods when nonlinear pricing takes the form of quantity discounts. When this condition does not strictly hold, the accuracy of our method will depend on the proportion of observations for which multiple brands are jointly purchased.

3. Proposed Model

We use a utility function that links the product class under consideration to an outside good. This allows us to model the trade-off in expenditure between the product class and other relevant products. More specifically, we use the Cobb-Douglas utility function

$$\ln u(x, z) = \alpha_0 + \alpha_x \ln u(x) + \alpha_z \ln(z), \qquad (4)$$

where $x = (x_1, \ldots, x_K)$ is the vector of the amount of each brand purchased, K represents the number of brands available in the product class, z represents the amount of the outside good purchased, and $u(x)$ denotes a subutility function. We consider two subutility structures with linear indifference curves to represent the near-perfect substitute nature of brands within a product class. The simplest structure is the linear utility model:

$$u(x) = \psi' x, \qquad (5)$$

where ψ_k is the marginal utility for Brand k, $\ln(\psi_k) = \nu_k + \varepsilon_k$, and ε_k is a stochastic element. This structure results in indifference curves that are linear and parallel. Alternatively, one could adopt the nonhomothetic structure of Allenby and Rossi (1991):

$$u(x) = \sum_{k=1}^{K} \psi_k(\bar{u}) x_k,$$
$$\ln \psi_k(\bar{u}) = \delta_k - \kappa_k \bar{u}(x, z) + \varepsilon_k, \qquad (6)$$

in which the deterministic part of marginal utility, $\bar{u}$, is implicitly defined. We note that Equation (6) differs from the nonhomothetic structure of Allenby and Rossi (1991) by including the outside good z. If the coefficient vector κ is strictly positive, this implicitly defined utility function results in linear indifference curves that fan out but do not intersect in the positive orthant. The ratio of the expectation of the marginal utility of two brands, $E[\psi_i(\bar{u})]/E[\psi_j(\bar{u})] =$

$\exp[(\delta_i - \delta_j) - \bar{u}(x,z)(\kappa_i - \kappa_j)]$ indicates that if $\kappa_i < \kappa_j$, then Brand i is relatively superior to Brand j. As utility increases, the ratio of marginal utility for Brand i increases relative to Brand j. Thus, the κ parameters govern the relative rates of rotation of the indifference curses. Equation (6) is therefore a generalization of Equation (5) in which the marginal utility is related to expenditure. In a standard choice model, both the linear and nonhomothetic utility specification results in corner solutions in which only one of the available brands is selected.

The consumer determines his/her expenditure on the K brands and the outside good by maximizing the utility function in (4) subject to the budget constraint

$$\sum_{k=1}^{K} p_k(x_k) + z = T, \tag{7}$$

where $p_k(x_k)$ is the price of x_k units of Brand k, and the price of the outside good is one. The price function $p_k(x_k)$ allows for the possibility of price discounts. Note that, if $p_k(x_k) = q_k x_k$ for each k, where q_k is a constant, then the budget constraint reduces to the standard linear budget constraint.

In contrast to approaches that rely on first-order conditions to relate model parameters to the data, our approach is to directly evaluate the above expression at all feasible solution points. At first glance, this appears to be a formidable task because of the stochastic elements in the utility specification. However, if the utility maximizing solution is restricted to corner solutions, in which only one element of x is nonzero, we obtain a solution that is easy to evaluate. We first outline our solution strategy and resulting likelihood, and then prove our proposition that the utility maximizing solution to the Cobb-Douglas function (Equation (4)) coupled with either linear subutility (Equation (5)) or nonhomothetic subutility (Equation (6)) subject to the budget constraint (Equation (7)) is at a corner.

Solution Procedure and Likelihood Specification

The likelihood function is derived from assumptions about stochastic elements in the utility function. In the standard random utility model (McFadden 1986) leading to the logit or probit likelihood, the vector of log-marginal utility ($\ln\psi$) is assumed to be stochastic with an additive error, $\ln\psi_k = \nu_k + \varepsilon_k$. Assumptions about the random utility error are used to derive the likelihood of observed demand. In the case of a linear utility model, $U(x) = \psi' x$, discrete choice probabilities are derived from the utility maximizing solution (choose x_k if ψ_k/p_k is maximum) that links observed demand to the latent utility parameters.

In our model of demand, the likelihood for the observed demand is derived for offerings in the product category $x = (x_1, x_2, \ldots, x_K)'$ and the outside good (z) from distributional assumptions of random utility. We note that the budget restriction in Equation (7) imposes an "adding-up" constraint that induces a singularity in the distribution of observed, utility maximizing demand (x, z) (see Kim et al. 2002). Therefore, only K error terms—one for each brand—are needed to derive the likelihood of observed demand, and we substitute $z = T - \sum_k p_k(x_k)$ for the outside good. We follow standard convention by assuming that the vector of log-marginal utility is stochastic with an additive error.

The solution strategy when only one element of x is nonzero involves two steps. In the first step, the optimal product quantity is determined for each brand separately. For the linear utility structure with $\ln\psi_k = \nu_k + \varepsilon_k$, the maximization occurs over all possible combinations of the available pack sizes, including multiple packs, (a):

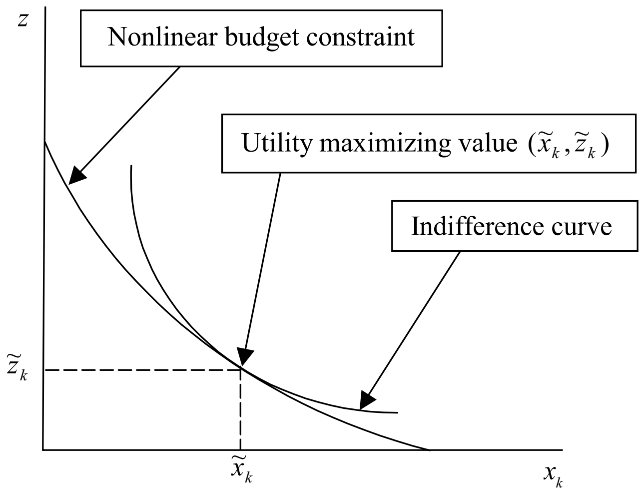
$$\begin{aligned} &\operatorname*{Max}_a \{\alpha_0 + \alpha_x \ln\psi_k x_{ka} + \alpha_z \ln(T - p_k(x_{ka}))\} \\ &\quad = \operatorname*{Max}_a \{\alpha_x \ln x_{ka} + \alpha_z \ln(T - p_k(x_{ka}))\}, \end{aligned} \tag{8}$$

where T is the budgetary allotment for the product and outside good, and $p_k(x_{ka})$ denotes the price of Brand k with packsize x_{ka}. We note that, if a retailer offers only a limited selection of package sizes, the consumer can consider alternative combinations of the available offerings. For example, if a retailer only offers six-packs of a beverage, the consumer can consider 6-, 12-, 18-, 24-, ... package bundles. The stochastic element, ε_k, is the same for each of the package bundles in Equation (8) and cancels from the expression. That is, because the same good is contained in each of the possible package bundles, the determination of the utility maximizing quantity is deterministic. Furthermore, by substituting the expression $T - p_k(x_{ka})$ for z, we ensure that the evaluated solution points correspond to the budget restriction.

When the available quantities are continuous rather than discrete, as in the case of the purchase of telephone services, the use of a grid search procedure will work well because it is one dimensional. As shown in Figure 2, the indifference curves associated with the Cobb-Douglas function do not intersect the axes, and, in principal, first-order conditions can be used to identify the point of tangency between the indifference curve and the budget line in the (x_k, z) plane. However, when the price schedule is piecewise linear, with price discounts beginning at specific values,

$$p_k(x_k) = \begin{cases} p_{k,M_k+1} x_k = p_{k,Low} x_k & \text{for } x_k \geq x_{k,Cut_{M_k}} \\ p_{kM_k} x_k & \text{for } x_{k,Cut_{M_k}-1} \leq x_k < x_{k,Cut_{M_k}} \\ \vdots & \\ p_{k1} x_k = p_{k,Hi} x_k & \text{for } x_k < x_{k,Cut_1}, \end{cases} \tag{9}$$

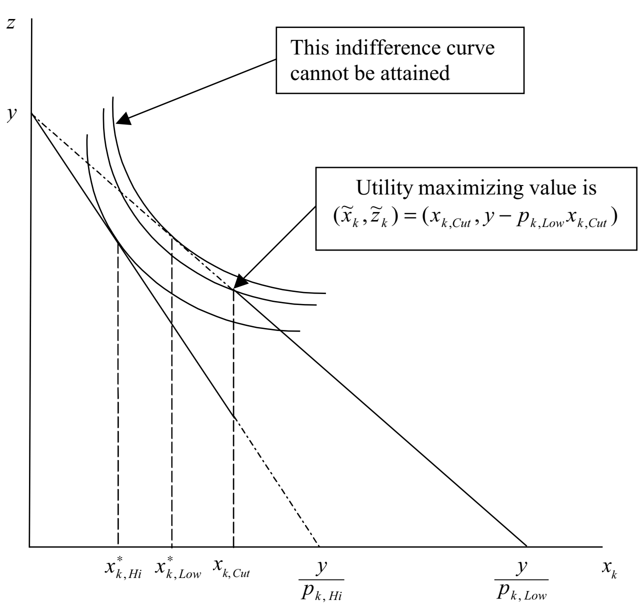
Figure 2 Utility Maximization in the (x, z) Plane for Continuous Nonlinear Prices



then the utility maximizing point in the (x_k, z) may be at the point of discontinuity of the budget line, and the identification becomes more complicated. Figure 3 illustrates that the utility maximizing value need not be at a point of tangency, but can be located at a point of discontinuity, in a price schedule. When using a grid search procedure to identify the point of utility maximization, it is therefore important to include the points of discontinuity in the grid.

A limitation of using a linear utility structure $u(x) = \psi' x$ is that the optimal quantity for each brand depends only on the prices for the brand, not on the quality of the brand (ν_k). The terms α_0 and $\ln \psi$ are constant over the alternative bundles in Equation (8), shifting the intercept of utility but not affecting the

Figure 3 Utility Maximization in the (x, z) Plane for a Price Schedule



optimal allocation of the budget between alternatives in the product class (x) and the outside good (z). This limitation of the Cobb-Douglas utility function can be overcome by using the nonhomothetic function in Equation (6). For the nonhomothetic function, the optimal quantity for each brand proceeds by solving for the package size that maximizes the expression:

$$\operatorname*{Max}_a \{\alpha_0 + \alpha_x \ln \psi_k(\bar{u}) x_{ka} + \alpha_z \ln(T - p_k(x_{ka}))\}$$
$$= \operatorname*{Max}_a \{\alpha_x \ln \psi_k(\bar{u}) + \alpha_x \ln x_{ka} + \alpha_z \ln(T - p_k(x_{ka}))\}, \tag{10}$$

where $\ln \psi_k(\bar{u}) = \nu_k + \varepsilon_k = \delta_k - \kappa_k \bar{u}(x, z) + \varepsilon_k$. As with the linear utility structure, determination of the optimal quantity for Brand k is deterministic because ε_k is the same for each package size. The quantity depends on the brand because κ_k does not cancel from the evaluation. The optimal quantity of the product demanded therefore depends on the brand under consideration.

The first step of the solution procedure, outlined here, identifies the optimal quantity given that Brand k is purchased. The second step of our method involves determining the probability that Brand k is purchased. In this second step, the stochastic element of marginal utility takes on a traditional role of generating a choice probability. We have:

$$\operatorname*{Max}_{ka} \{\ln u(x_{ka}, T - p_k(x_{ka}))\}$$
$$= \operatorname*{Max}_k \Big[\operatorname*{Max}_{a|k} \{\ln u(x_{ka}, T - p_k(x_{ka}))\}\Big]$$
$$= \operatorname*{Max}_k \big[\alpha_0 + \alpha_x \ln u(x^*_k) + \alpha_z \ln(T - p_k(x^*_k))\big], \tag{11}$$

where x^*_k indicates the optimal quantity for Brand k that is identified in the first step. Substituting the linear subutility expressions in Equation (10) results in the expression:

$$= \operatorname*{Max}_k \big[\alpha_0 + \alpha_x(\nu_k + \varepsilon_k) + \alpha_x \ln(x^*_k)$$
$$+ \alpha_z \ln(T - p(x^*_k))\big], \tag{12}$$

and assuming that $\varepsilon_k \sim \mathrm{EV}(0, 1)$ results in the choice probability:

$$\Pr(x_i) = \frac{\exp[\nu_i + \ln(x_i) + (\alpha_z/\alpha_x)\ln(T - p_i(x_i))]}{\sum_{k=1}^{K} \exp[\nu_k + \ln(x^*_k) + (\alpha_z/\alpha_x)\ln(T - p_k(x^*_k))]}, \tag{13}$$

where x_i is the observed demand. In the estimation procedure, the observed x_i is used for the selected brand, whereas x^*_k is used for the brands not selected. Alternatively, assuming nonhomothetic subutility in Equation (10) results in the expression:

$$\Pr(x_i) = \frac{\exp[\delta_i - \kappa_i \bar{u}^i + \ln(x_i) + (\alpha_z/\alpha_x)\ln(T - p_i(x_i))]}{\sum_{k=1}^{K} \exp[\delta_k - \kappa_k \bar{u}^k + \ln(x^*_k) + (\alpha_z/\alpha_x)\ln(T - p_k(x^*_k))]}, \tag{14}$$

where $\bar{u}^i = \bar{u}(x_i^*, T - p_i(x_i^*))$ is the deterministic part of the utility derived from the consumption of x_i^* units of Brand i, with the remainder of the budget allocation, $T - p_i(x_i^*)$, devoted to the outside good. From Equation (6), we see that $\bar{u}^i$ solves the equation:

$$\ln \bar{u}^i = \delta_i - \kappa_i \bar{u}^i + \ln(x_i^*) + (\alpha_z/\alpha_x)\ln(T - p_i(x_i^*)). \quad (15)$$

Details of the properties of the nonhomothetic function and its estimation can be found in Allenby and Rossi (1991).

We note that not all parameters of the Cobb-Douglas specification are theoretically identified. In particular, because an arbitrary rescaling of utility by a constant and multiplicative factor can represent the same preference ordering, we set α_0 to zero, α_x to one, and estimate $\alpha_z^* = \alpha_z/\alpha_x$ subject to the constraint that it takes on a positive value.

This strategy is dependent on conducting an initial search along each of the axes of demand. We next provide a proof that corner solutions are consistent with utility maximization when indifference curves are linear and prices are either constant or monotonically decreasing with quantity.

Proof of Corner Solution

Assume that the amount of expenditure for the product is fixed at T^*. The proof holds for any value of T^*, and this assumption therefore does not restrict the generality of the proof. For any continuous convex budget set defined as $\sum_k p(x_k) < T^*$, there exists a less restrictive linear set $\sum_k \hat{p}_k x_k < T^*$ that contains the convex set, where $\hat{p}_k = T^*/x_k^*$ is the per-unit price of allocating all the expenditure to Brand k (i.e., x_k^* s.t. $p(x_k^*) = T^*$). As shown in Figure 4, the dominating linear budget set is equal to the convex budget set at each axis and is greater than the convex set at all interior points. For the price schedule with discounts at specific quantities (Equation (9)), a dominating linear budget set also exists that is equal to the actual set at each axis and is greater than or equal to the actual set in the interior. Figure 5 illustrates such a budget set when one price cut exists. Because the dominating linear budget set is less restrictive than the convex set, the maximum utility in the linear set is greater than or equal to the maximum utility in the convex set. We note that the utility maximizing solution for the linear budget set is a corner solution when the indifference curves are linear (e.g., Equations (8) and (10)). Therefore, the utility maximizing solution for any convex budget set contained by the linear set must also be at the same corner.

Figure 4 Continuous Convex Budget Set and Dominating Linear Set

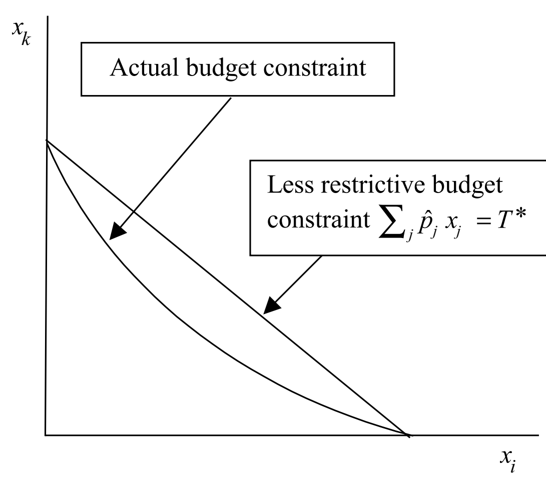
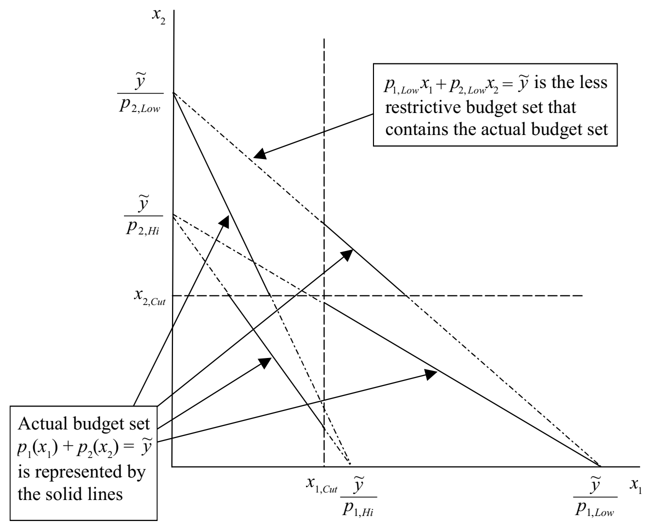


Figure 5 Piecewise Linear Budget Set and Dominating Linear Set



4. Data

Our model is illustrated with a scanner panel data of light-beer purchases. We study the three dominant brands in the domestic product category—Miller Lite, Bud Light, and Coors Light. The data set is comprised of 20,914 purchase occasions for 2,282 households making beer purchases at grocery stores. Table 2 reports the packages sizes and package types under study. Our analysis includes beer packaged in nonreturnable bottles and cans, for which there are 84 different varieties for the three brands. Table 2 also reports the choice shares, average prices, and the frequency of merchandising activity for each of the choice alternatives.

Inspection of Table 2 reveals the following data characteristics. Bud Light has the highest overall choice share at 0.43, followed by Miller Lite at 0.35 and Coors Light at 0.22. The choice shares for the smaller package sizes are approximately equal, whereas the choice share for larger package sizes (e.g., 12-packs) favors Bud Light. For example, the choice share of Miller Lite 6-packs of 12 oz. bottles and cans (Items 14 and 15) is equal to 0.0494, Bud

Table 2 Description of the Data

Item ID	Brand ID	Oz./Unit	No. of Units/Pack	Package Form	Volume/ Pack	Price/ Oz. ($)	Display Frequency	Feature Frequency	Choice Share	Average Unit Price ($) by Trademark	Average Choice Share by Trademark
1	1	7	1	2	7	0.0967	0.0000	0.0000	0.0000	0.0536	0.3504
2	1	12	1	2	12	0.0649	0.0000	0.0000	0.0014		
3	1	16	1	2	16	0.0564	0.0000	0.0000	0.0002		
4	1	22	1	2	22	0.0483	0.0000	0.0000	0.0019		
5	1	24	1	1	24	0.0482	0.0004	0.0007	0.0003		
6	1	32	1	2	32	0.0495	0.0001	0.0000	0.0041		
7	1	32	1	1	32	0.0393	0.0000	0.0004	0.0031		
8	1	40	1	2	40	0.0511	0.0000	0.0000	0.0002		
9	1	7	6	2	42	0.0706	0.0000	0.0000	0.0009		
10	1	8	6	1	48	0.0550	0.0000	0.0000	0.0000		
11	1	10	6	1	60	0.0641	0.0000	0.0000	0.0001		
12	1	16	4	2	64	0.0544	0.0000	0.0000	0.0008		
13	1	16	4	1	64	0.0526	0.0000	0.0000	0.0000		
14	1	12	6	1	72	0.0587	0.0011	0.0011	0.0166		
15	1	12	6	2	72	0.0588	0.0028	0.0007	0.0328		
16	1	16	6	1	96	0.0525	0.0014	0.0002	0.0036		
17	1	10	12	1	120	0.0603	0.0016	0.0000	0.0000		
18	1	12	12	1	144	0.0538	0.1355	0.1016	0.0951		
19	1	12	12	2	144	0.0541	0.0897	0.0743	0.0553		
20	1	12	15	1	180	0.0420	0.0242	0.0070	0.0085		
21	1	16	12	1	192	0.0428	0.0015	0.0003	0.0002		
22	1	12	18	1	216	0.0497	0.0735	0.0281	0.0272		
23	1	12	18	2	216	0.0475	0.0005	0.0000	0.0000		
24	1	10	24	1	240	0.0559	0.0039	0.0006	0.0000		
25	1	12	24	1	288	0.0493	0.1318	0.0784	0.0788		
26	1	12	24	2	288	0.0485	0.0001	0.0000	0.0002		
27	1	12	30	1	360	0.0396	0.0922	0.0305	0.0191		
28	1	16	24	2	384	0.0427	0.0000	0.0000	0.0000		
29	1	16	24	1	384	0.0474	0.0000	0.0000	0.0000		
30	2	12	1	2	12	0.0591	0.0000	0.0000	0.0006	0.0548	0.4327
31	2	16	1	2	16	0.0610	0.0000	0.0000	0.0000		
32	2	22	1	2	22	0.0521	0.0000	0.0000	0.0044		
33	2	24	1	1	24	0.0539	0.0008	0.0002	0.0052		
34	2	32	1	2	32	0.0565	0.0000	0.0000	0.0040		
35	2	40	1	2	40	0.0507	0.0009	0.0000	0.0020		
36	2	7	6	2	42	0.0724	0.0001	0.0000	0.0051		
37	2	8	6	2	48	0.0632	0.0000	0.0000	0.0005		
38	2	7	8	2	56	0.0744	0.0000	0.0000	0.0000		
39	2	10	6	1	60	0.0642	0.0002	0.0000	0.0000		
40	2	16	4	2	64	0.0561	0.0029	0.0000	0.0011		
41	2	8	8	1	64	0.0541	0.0000	0.0000	0.0011		
42	2	12	6	1	72	0.0591	0.0015	0.0015	0.0302		
43	2	12	6	2	72	0.0594	0.0031	0.0004	0.0328		
44	2	16	6	1	96	0.0536	0.0121	0.0000	0.0048		
45	2	16	6	2	96	0.0549	0.0000	0.0000	0.0000		
46	2	10	12	1	120	0.0633	0.0009	0.0001	0.0002		

Light (Items 42 and 43) is equal to 0.0630, and Coors Light (Items 72 and 73) is 0.0521, whereas the choice shares of 12-packs is equal to 0.1504 for Miller Lite (Items 18 and 19), 0.2188 for Bud Light (Items 47 and 48), and 0.1280 for Coors Light (Items 77 and 78). Data also indicate that bottles are somewhat more preferred to cans for 6-packs (e.g., Items 14 vs. 15, 42 vs. 43, and 72 vs. 73), but that cans are preferred to bottles when consumers purchase larger quantities of beer (e.g., Items 18 vs. 19, 47 vs. 48, and 77 vs. 78). These results indicate that the nonhomothetic specification (Equation (6)) may fit the data better because the marginal utility for the brands depends on the level of expenditure. However, the shift in brand and package preferences could also be attributed to differences in consumer tastes (e.g., households that consume large quantities of beer may simply prefer Bud Light). Therefore, we fit both the linear subutility model (Equation (5)) and the nonhomothetic subutility model (Equation (6)).

Table 2 (Cont'd.)

Item ID	Brand ID	Oz./Unit	No. of Units/Pack	Package Form	Volume/ Pack	Price/ Oz. ($)	Display Frequency	Feature Frequency	Choice Share	Average Unit Price ($) by Trademark	Average Choice Share by Trademark
47	2	12	12	1	144	0.0541	0.1984	0.1236	0.1589		
48	2	12	12	2	144	0.0547	0.1392	0.0696	0.0599		
49	2	7	24	2	168	0.0597	0.0000	0.0000	0.0000		
50	2	12	15	1	180	0.0431	0.0083	0.0018	0.0065		
51	2	12	18	1	216	0.0512	0.0989	0.0310	0.0273		
52	2	10	24	1	240	0.0567	0.0032	0.0006	0.0011		
53	2	12	20	2	240	0.0441	0.0138	0.0018	0.0024		
54	2	12	24	1	288	0.0502	0.2091	0.0948	0.0740		
55	2	12	24	2	288	0.0475	0.0044	0.0008	0.0005		
56	2	24	12	1	288	0.0439	0.0005	0.0000	0.0000		
57	2	22	15	2	330	0.0438	0.0000	0.0000	0.0000		
58	2	12	30	1	360	0.0405	0.0390	0.0164	0.0100		
59	2	16	24	1	384	0.0468	0.0000	0.0000	0.0000		
60	3	7	1	2	7	0.0732	0.0000	0.0000	0.0000	0.0593	0.2170
61	3	12	1	2	12	0.0604	0.0000	0.0000	0.0007		
62	3	16	1	2	16	0.0609	0.0000	0.0000	0.0003		
63	3	18	1	2	18	0.0732	0.0018	0.0000	0.0001		
64	3	22	1	2	22	0.0496	0.0000	0.0000	0.0006		
65	3	24	1	1	24	0.0501	0.0002	0.0000	0.0007		
66	3	32	1	2	32	0.0557	0.0000	0.0000	0.0019		
67	3	40	1	2	40	0.0511	0.0000	0.0000	0.0011		
68	3	7	6	2	42	0.0712	0.0000	0.0000	0.0019		
69	3	8	6	1	48	0.0904	0.0000	0.0000	0.0000		
70	3	10	6	1	60	0.0646	0.0000	0.0000	0.0000		
71	3	16	4	2	64	0.0597	0.0003	0.0000	0.0009		
72	3	12	6	1	72	0.0589	0.0043	0.0026	0.0294		
73	3	12	6	2	72	0.0593	0.0008	0.0002	0.0227		
74	3	18	4	2	72	0.0748	0.0125	0.0000	0.0006		
75	3	16	6	1	96	0.0533	0.0001	0.0001	0.0007		
76	3	10	12	1	120	0.0633	0.0006	0.0000	0.0004		
77	3	12	12	1	144	0.0540	0.1255	0.1041	0.0966		
78	3	12	12	2	144	0.0540	0.0575	0.0460	0.0314		
79	3	7	24	2	168	0.0773	0.0000	0.0000	0.0000		
80	3	12	18	1	216	0.0499	0.0695	0.0243	0.0183		
81	3	18	12	2	216	0.0358	0.0000	0.0000	0.0000		
82	3	10	24	1	240	0.0560	0.0018	0.0005	0.0001		
83	3	12	24	2	288	0.0477	0.0003	0.0000	0.0000		
84	3	12	30	1	360	0.0392	0.0395	0.0165	0.0085		

Note:

Variable	Description
Brand ID	1 = Miller Lite, 2 = Bud Light, 3 = Coors Light
Oz./unit	Oz. per bottle or can
No. of units/pack	No. of units of cans or bottles included in the pack
Package form	1 = cans, 2 = nonreturnable bottles
Volume/pack	Total volume contained in the pack

We note that Budweiser engages in significantly more display and feature activity than either Miller or Coors. A summation of the display and feature frequencies reported in Table 2 reveals that some package of Bud Light is on display in 74% of the purchases, compared with 56% for Miller Lite and 31% for Coors Light. Feature activity is approximately equal for Bud Light and Miller Lite (33% of the time), but lower for Coors Light (19%). An interesting issue to investigate is whether the superior performance of Bud Light in large package sizes is because of it being a superior product offering, as measured by the nonhomothetic coefficients in Equation (6), versus engaging in more frequent merchandising activity.

There exists large intrahousehold variation in grocery expenditures, with many shopping trips totaling more than $200, whereas others are much less. Across these purchase occasions, the nonbeer expenditures take on a varied meaning. For small shopping trips, the outside good includes snacks and a few miscellaneous items, whereas for the larger shopping trips the outside good is comprised of a much broader array of items. Rather than equating the observed grocery expenditure to the budgetary allotment (T) in Equations (13) and (14), we treat the budgetary allotment as an unknown parameter and estimate it from the data. In our analysis, we specify a prior distribution for T and derive posterior estimates. The prior specification reflects managerial judgment about the amount of money consumers allocate to beer and substitutable goods. Markov chain Monte Carlo for estimating the budgetary allotment and other model parameters are provided in the Appendix.

5. Results

The linear (Equation (5)) and nonhomothetic (Equation (6)) subutility structures were imbedded to the Cobb-Douglas function (Equation (4)) and estimated as hierarchical Bayes models (Gelfand and Smith 1990). As discussed previously, some parameters require restriction to ensure algebraic signs that conform to economic theory. In the Cobb-Douglas model, we set $\alpha_x = 1$ and $\alpha_z = \exp(\alpha_z^*)$ and estimate α_z^* unrestricted. In the linear subutility function, we follow standard convention and set one of the brand intercepts (Miller Lite) to zero. In the nonhomothetic subutility function, we restrict κ to be positive by estimating $\kappa^* = \ln \kappa$ with κ^* unrestricted, and set κ^* for Miller Lite to zero.

The deterministic portion of log marginal utility for the linear model is specified as:

$$\nu_i = \beta_{0i} + \beta_b(\text{bottle}_i) + \beta_d(\text{display}_i) + \beta_f(\text{feature}_i), \quad (16)$$

where β_{0i} is the preference coefficient for Brand i. The covariates "bottle," "display," and "feature" are coded as dummy variables for each brand. This specification assumes that the marginal utility of a brand can be influenced by the package type and merchandising activity of the retailer. The nonhomothetic model specification is:

$$\begin{aligned} \nu_i &= \delta_i - \kappa_i \bar{u}(x, z); \\ \delta_i &= \beta_{0i} + \beta_d(\text{display}_i) + \beta_f(\text{feature}_i); \qquad (17) \\ \kappa_i &= \exp[\kappa_{0i} + \beta_b(\text{bottle}_i)]. \end{aligned}$$

As reported in Allenby and Rossi (1991), the β_{0i} intercept parameters are redundant if the κ_i parameters take on nonzero values, and are therefore set to zero. Alternative model specifications were investigated (e.g., bottle$_i$ related to δ_i instead of κ_i), but did not result in better fit relative to Equation (15). Household heterogeneity was allowed for all parameters $\theta = (\kappa', \alpha, \beta')'$ and specified as a multivariate normal distribution:

$$\theta \sim \text{Normal}(\bar{\theta}, V). \quad (18)$$

Table 3 Model Fit

Subutility Model	Log-Marginal Density
Linear	−4864.99
Nonhomothetic	−4859.79

A total of 50,000 iterations of the Markov chain were executed, with the last 25,000 iterations used to estimate model parameters. Convergence was checked by starting the chain from multiple start points and noting common convergence, and through inspection of time-series plots.

Table 3 reports the log-marginal density of the linear and nonhomothetic models. The marginal density was estimated using the importance sampling procedure of Newton and Raftery (1994). The marginal densities are high and translate to an in-sample hit probability of approximately 0.75 across the 20,914 observations. A reason for the exceptional model fit, in the presence of more than 80 choice alternatives, is that there are only three error terms in the model—one for each brand.

The model fit statistics indicate that the nonhomothetic subutility model fits the data better than the linear subutility model. This implies that the shift in brand preference to Bud Light, and package preference to bottles, reported in Table 2 cannot be solely attributed to differences in household tastes. Bottles weigh more than cans, which may partially explain the shift in preference to cans during grocery trips with larger beer expenditures. Bud Light is the market share leader and has a less bitter composition than Miller Lite. Having a less bitter composition decreases the likelihood of satiation when larger quantities are consumed.

Table 4 reports aggregate parameter estimates for the nonhomothetic subutility model. Bud Light has the smallest aggregate estimate of κ, indicating that at higher levels of utility ($u(x, z)$), it is preferred to Miller Lite and Coors Light. Coors Light has the largest aggregate estimate of κ, indicating that it is the least-preferred brand when total utility is high. The aggregate estimate of $\alpha_z^* = \ln \alpha_z = -0.548$ translates to an estimate of α_z equal to 0.578. Recall that we have restricted our analysis to household

Table 4 Aggregate Coefficient Estimates (Posterior Standard Deviations)

Parameter	Estimate
Bud Light κ^*	−1.230 (0.228)
Coors Light κ^*	3.006 (0.322)
$\alpha^* = \ln(\alpha_z/\alpha_x)$	−0.548 (0.068)
Bottle	0.233 (0.032)
Feature	1.026 (0.131)
Display	0.766 (0.090)

Table 5 Covariance/*Correlation* Matrix of Random Effects (Posterior Standard Deviations)

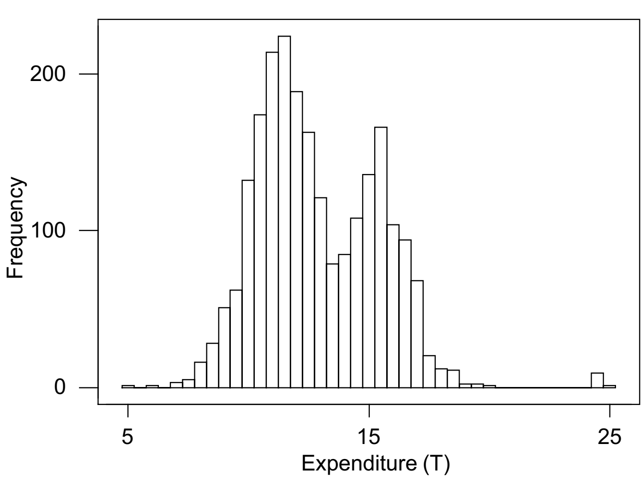
Bud Light κ^*	72.52 (6.74)	*−0.03*	*−0.02*	*−0.03*	*−0.14*	*0.04*
Coors Light κ^*	−2.34 (2.60)	84.38 (4.01)	*0.03*	*0.02*	*0.29*	*0.10*
α^*	−0.11 (0.70)	0.18 (0.60)	0.49 (0.05)	*0.00*	*0.03*	*0.07*
Bottle	−0.11 (0.32)	0.08 (0.32)	0.00 (0.03)	0.24 (0.02)	*0.02*	*−0.03*
Feature	−1.13 (1.72)	2.46 (1.57)	0.02 (0.05)	0.01 (0.03)	0.85 (0.18)	*0.01*
Display	0.25 (0.89)	0.72 (0.80)	0.04 (0.04)	−0.01 (0.02)	0.01 (0.06)	0.61 (0.07)

shopping trips in which beer is purchased, which accounts for the large amount of utility derived from the beer category ($\alpha_x = 1.00$) relative to the outside good ($\alpha_z = 0.578$). We discuss the implications of the Cobb-Douglas estimates here. On average, we find that households prefer cans when the grocery trip involves large expenditures, and bottles when the total expenditure is smaller. Finally, the feature and display coefficients have positive algebraic signs, as expected, with the effect of feature advertisement larger than the effect of a display.

The covariance matrix of the distribution of heterogeneity is reported in Table 5. Variances and covariances are reported in the lower left of the table, and the associated correlations are reported in the upper right portion of the table. The diagonal entries for the brands are large, indicating that there is a wide dispersion of preferences for the brands. In addition, we find that the off-diagonal entries are small in magnitude, indicating that brand preferences are not strongly associated with feature and display sensitivity, preference for bottle versus can, or preference for the outside good, z.

Posterior estimates of household expenditures (T) are displayed in Figure 6. The mean of the distribution is equal to \$12.86, the interquartile range is (\$10.97, \$14.98), and the distribution is bimodal. The correlation between the posterior estimates and the Cobb-Douglas parameter α_z^* is approximately zero, implying that there is little evidence in the data of an association between the amount of budget allocation and price sensitivity. Hence, retailers and manufacturers have little ability to price discriminate between high-volume and low-volume users. Additional implications of the parameter estimates are explored later.

Figure 6 Posterior Estimates of Household Expenditure (T)



6. Implications

We investigate three issues related to pricing and product assortment. The first is the effect of a price reduction for a package size (12-pack) for a particular brand (Miller Lite) in terms of substitution of demand across brands and with the outside good. Of particular interest is the extent to which the increase in demand caused by a price reduction can be attributed to brand switching versus an increase in primary demand in the product category. The second issue relates to the expected demand for a particular package size. Manufacturers often speculate about the existence of price points, or threshold values at which individual households become willing to consider a given package size. Our economic model is capable of identifying the price at which a particular package size becomes the most-preferred quantity for a household. Finally, our model provides a valid measure of utility that can be used to assess the change in consumer welfare by computing the compensating value associated with changes in the assortment of package sizes. We use the model to measure the utility consumers derive from the various package sizes by considering the impact of removing alternative package sizes from the current mix of offerings.

The derivation of model implications is a function of the model parameters, the available offerings, and current pricing policies of the retailers. In the analysis presented here, we do not attempt to explore supply-side implications of our model by identifying the optimal number and size of a brand's offerings (see Villas-Boas and Winer 1999, Chintagunta 2000). In addition, because our analysis conditions on light-beer purchases, the pricing implications do not investigate issues, such as purchase timing, stockpiling, and substitution from the regular beer category. We leave these issues as extensions for future work that extend our basic model.

Table 6 Expected Demand

Brand	Current Prices	Miller 12-Pack on Sale (20% Off)
Miller Lite	67.41 oz.	77.95 oz.
Bud Light	75.72 oz.	75.64 oz.
Coors Light	34.24 oz.	34.19 oz.
Outside good	5.46 units	5.36 units

The Effect of a Price Reduction on Primary and Secondary Demand

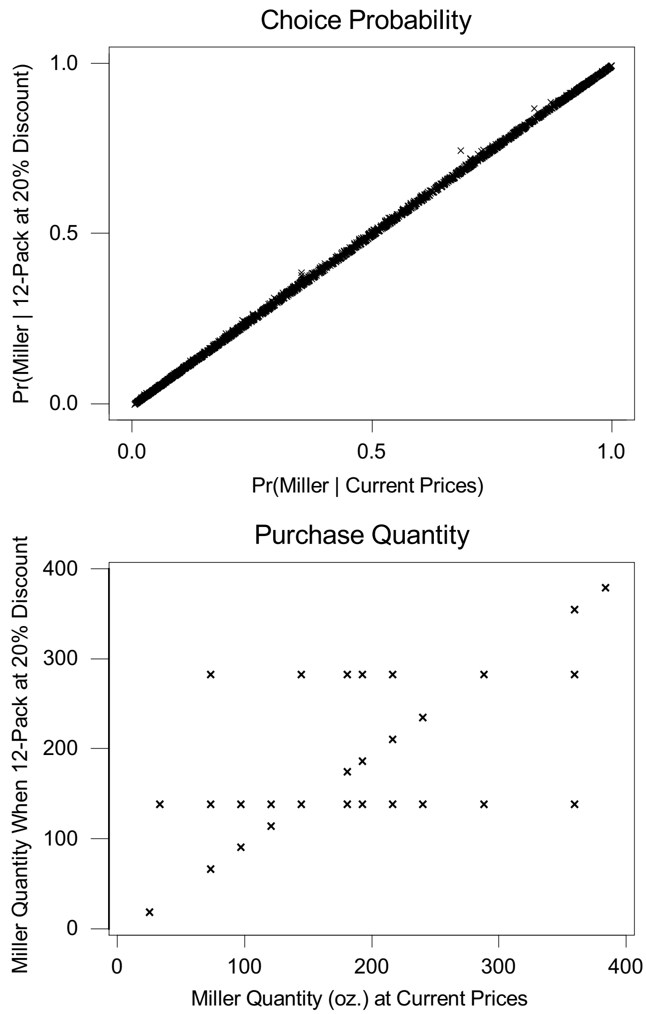
Table 6 reports the expected demand for the three brands and the outside good under two conditions: (1) when prices are distributed as in the data set; and (2) when the Miller 12-pack items (Items 18 and 19 in Table 2) are reduced by 20%. Expected demand for a brand is computed as the product of the optimal quantity for each brand, x_i^*, and the probability of purchasing this quantity $\Pr(x_i^*)$ in Equation (13), integrated over the empirical distribution of the explanatory data.

The estimates of expected demand indicate that a price reduction by Miller Lite has little effect on the demand for Bud Light and Coors Light. The demand for Miller Lite increases from 67.41 oz. per purchase to 77.95 oz. per purchase, whereas demand for the other brands is nearly constant. Figure 7 decomposes the change in demand for Miller Lite into the change in the brand-choice probability and the change in the expected purchase quantity for each observation in the data set. The top panel shows that the choice probability does not change and is nearly identical for the regular and discounted prices. The purchase quantity, however, is substantially different, with demand shifting to the two horizontal lines at 144 oz. (a 12-pack of 12-oz. bottles) and 288 oz. (two 12-packs). At the reduced price, the increase in demand is obtained from Miller drinkers who allocate greater expenditure from the outside good to the product category (i.e., an increase in primary demand), with the brand-choice probabilities nearly unchanged (i.e., no change in secondary demand).

Price Points

Changes in price result in a redistribution of the quantity demand. The demand for a specific quantity (package size) is a function of the prices of the other alternatives and household parameters. Our model permits calculation of the price at which a particular package size is most preferred by a household by identifying the highest price at which a particular quantity (x^*) yields maximum utility in Equation (8) or (10). Figure 8 displays the cumulative distribution of prices at which the 12-pack of Miller Lite is the most preferred Miller offering. When priced at \$10, the 12-pack is the most-preferred package size for 2% of the population, and at \$5 it is the most preferred package size for 49% of the population. Even at very low prices, not everyone is willing to purchase a 12-pack. The reason is that many households are estimated to have high utility for the other brands. When coupled with low utility for the outside good (α_z), it is not possible to reduce the price of Miller Lite low enough to induce these individuals to purchase a 12-pack.

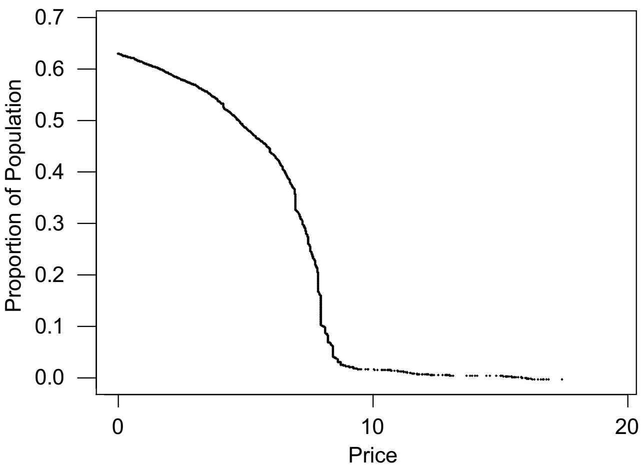
Figure 7 Effect of Price Reduction on Choice Probability and Quantity Demanded



Utility of an Assortment

A common problem in retailing is determining the variety provided by alternative packages and variants of an offering. The depth and breadth of an assortment is not well reflected by the total number of stock keeping units (SKUs) in a product category because many of the offerings are nearly identical to each other. Because our model is derived from an economic model, we can use utility as a scalar measure in assessing the value of specific offerings. That is, we can compute the average consumer utility for the current set of offerings across the purchases and

Figure 8 **Cumulative Distribution of 12-Pack Price Points for Miller Lite**



investigate the decrease in utility as specific offerings are removed. These differences in utility can be translated into a dollar metric by computing the compensating value, defined as the increase in the budget (T) needed to return to the original level of utility.

Table 7 reports on the change in utility and associated compensating value of removing alternative package sizes across all brands. The current utility (per purchase) of 3.35, for example, declines to 3.05 when 12-packs are no longer made available to consumers. This translates to a compensating value of 30 cents per purchase, indicating a substantial decrease in consumer utility. Our analysis indicates that consumers would be least affected by the removal of six-packs from the shelves.

7. Conclusion

The analysis of demand for packaged goods requires models that reflect both the brand and the quantity contained in the package. The discreteness of package sizes and the nonlinearity of package prices produce complexities in the analysis of demand data with economic models. Standard first-order conditions and identities used with continuous quantity and fixed price models do not hold, resulting in the need for models of utility maximization that involve the direct evaluation of the utility function. This evaluation is potentially difficult because of the stochastic elements of the demand function. However, we show that, when the brands under study are near-perfect substitutes and prices reflect quantity discounts, the evaluation is easy to perform.

Table 7 **The Utility of Alternative Assortments**

Assortment	Utility Per Purchase	Compensating Value ($)
Current offerings	3.3516	0.000
Remove 6-packs	3.3434	0.005
Remove 12-packs	3.0477	0.300
Remove 18-packs	3.3161	0.026
Remove 24-packs	3.3094	0.033

Our approach to deriving the likelihood function for the utility maximizing solution is to split the task of utility maximization into two steps. In the first step, the best package size is identified for each brand under study. The identification involves a deterministic search whenever error terms in the model are brand specific and not specific to brand-pack combinations. The error term cancels for the different package sizes. The second step involves a comparison of the brand-specific maximum utilities. The proposed approach can accommodate a variety of models, including linear and nonhomothetic subutility specifications, and more descriptive models of choice, such as ordered logit and probit models.

The advantages of using an economic specification are that the model becomes very parsimonious and the outside good (z) can be reasonably incorporated into the model specification. The use of either a multinomial or ordered-choice model to analyze the beer data would require an order of magnitude increase in the number of parameters to account for the many different package sizes. Furthermore, the outside good is required when deriving realistic pricing policies. Without its presence, the prices of all offerings can be raised, and demand for the product category remains constant. Substitution into and away from the category ensures that optimal prices are identified as not being too excessive. Although our Cobb-Douglas specification of utility for the product and outside good (Equation (10)) is somewhat restrictive when coupled with a linear subutility function, we argue that the use of a richer subutility specification for the offerings within the product class (e.g., a nonhomothetic specification) provides a flexible model structure that allows us to exploit the computational benefits of the two-step estimation approach.

Future research will allow us to examine the benefits of alternative utility functions and stochastic specifications. We derive our demand specification by assuming that the log-marginal utility for each brand is stochastic, and noting that the "adding-up" constraint imposes a singularity that alleviates the need to specify an additional stochastic element for the outside good. Alternative error specifications may be preferable. In particular, specifying the utility for the outside good as stochastic would lead to a solution procedure in which the first-stage maximization would not be deterministic. However, this would increase the complexity of computing estimates of utility function parameters, but may lead to a less restrictive specification in some instances.

Our model assumes that the price schedule is strictly concave with respect to quantity. There may be situations and instances when this is not true. For example, a retailer may deeply discount an intermediate package size (e.g., a 12-pack) such that the per-unit price is less than a larger package size (e.g., a 24-pack). In cases such as this, the utility maximizing solution need not exist at a corner, but instead may be associated with an interior point where consumers purchase more than one brand. The extent to which this creates a problem is directly revealed by the data by noting the proportion of occurrences where multiple brands are simultaneously purchased. If the proportion is large, then our proposed model is not valid for the data and a more flexible model specification is required, such as the nonlinear model of Kim et al. (2002) that leads to interior solutions.

Despite these limitations, the proposed model can serve as the kernel for many extensions. Examples include models of cross-category demand (Ainslie and Rossi 1998), purchase incidence (Chiang 1991, Chintagunta 1993, Arora et al. 1998), and the incorporation of supply-side issues (Yang et al., 2003). Discrete package sizes and nonlinear prices are commonly found in marketing and applied demand analysis. Our model of demand for near-perfect substitutes can accommodate these data characteristics, and is parsimonious, flexible, and easy to implement. We are, therefore, hopeful that many of these extensions will be realized.

Acknowledgments
Insightful comments were also provided by seminar participants at the University of California, Berkeley, University of Chicago, University of British Columbia, the University of Houston, University of Cincinnati, Purdue University, and the University of Wisconsin–Milwaukee.

Appendix: Markov Chain Monte Carlo Estimation
Estimation is carried out by sequentially generating draws from the following distributions:

1. Generate $\{\theta_h', T_h\} = \{\kappa_h', \alpha_h, \beta_h', T_h\}$ for $h = 1, \ldots, H$ households

$$\pi(\theta_h, T_h|^*) \propto \Pi_j \Pr(x_{ij} \mid \theta_h', T_h) \times \pi(\theta_h \mid \bar{\theta}, V) \times \pi(T_h \mid a, b)$$

$\Pr(x_{ij} \mid \theta_h', T_h)$ is the likelihood given by Equation (13) or (14), where "j" indexes the choice occasions and "i" denotes the brand selected on occasion j. The first-stage search in Equation (8) or (10) is conducted at each iteration of the Markov chain, for each observation, to arrive at the optimal quantity, x_i^*, for each choice alternative.

$\pi(\theta_h \mid \bar{\theta}, V)$ is the distribution of heterogeneity (Equation (16)).

$\pi(T_h \mid a, b)$ is the prior distribution for the budget limit, assumed normal.

$$a = 10$$
$$b = 100.$$

Draws of the conditional distribution are obtained with the Metropolis-Hastings algorithm with a random walk chain.

2. Generate $\bar{\theta}$

$$\pi(\bar{\theta} \mid \{\theta_h\}, V) = \text{Normal}\left(\textstyle\sum_h \theta_h / H, V/H\right)$$
$$H = \text{number of households} = 2{,}282.$$

3. Generate V

$$\begin{aligned}\pi(V \mid \{\theta_h\}, \bar{\theta}) &= \text{Inverted Wishart}\\ &\quad \cdot \left(g_0 + H, G_0 + \textstyle\sum_h (\theta_h - \bar{\theta})(\theta_h - \bar{\theta})'\right)\\ g_0 &= \text{prior degrees of freedom} = 50\\ G_0 &= \text{prior sum of squares and cross-products}\\ &= 50\text{I}.\end{aligned}$$

References

Ainslie, Andrew, Peter E. Rossi. 1998. Similarities in choice behavior across product categories. *Marketing Sci.* **17**(2) 91–106.

Allenby, Greg M. 1989. A unified approach to identifying, estimating and testing demand structures with aggregate scanner data. *Marketing Sci.* **8**(3) 265–280.

Allenby, Greg M., Peter J. Lenk. 1994. Modeling household purchase behavior with logistic normal regression. *J. Amer. Statist. Association* **49** 1218–1231.

Allenby, Greg M., Peter E. Rossi. 1991. Quality perceptions and asymmetric switching between brands. *Marketing Sci.* **10** 185–204.

Allenby, Greg M., Peter E. Rossi. 1999. Marketing models of consumer heterogeneity. *J. Econometrics* **89** 57–78.

Andrews, Rick L., Ajay K. Manrai. 1999. MDS maps for product attributes and market response: An application to scanner panel data. *Marketing Sci.* **18**(4) 584–604.

Arora, N., G. M. Allenby, J. L. Ginter. 1998. A hierarchical Bayes model of primary and secondary demand. *Marketing Sci.* **17** 29–44.

Bell, David R., James M. Lattin. 1998. Shopping behavior and consumer preference for store price format: Why "large basket" shoppers prefer EDLP. *Marketing Sci.* **17**(1) 66–88.

Blattberg, Robert C., Kenneth J. Wisniewski. 1989. Price-induced patterns of competition. *Marketing Sci.* **8**(4) 291–309.

Bolton, Ruth N. 1989. The relationship between market characteristics and promotional price elasticities. *Marketing Sci.* **8**(2) 153–169.

Bucklin, Randolph E., James M. Lattin. 1991. A two-state model of purchase incidence and brand choice. *Marketing Sci.* **10**(1) 24–39.

Chang, Kwangpil, S. Siddarth, Charles B. Weinberg. 1999. The impact of heterogeneity in purchase timing and price responsiveness on estimates of sticker shock effects. *Marketing Sci.* **18**(2) 178–192.

Chiang, J. 1991. A simultaneous approach to the whether, what and how much to buy questions. *Marketing Sci.* **10** 297–315.

Chiang, J. 1995. Competing coupon promotions and category sales. *Marketing Sci.* **14**(1) 105–122.

Chintagunta, Pradeep K. 1992. Estimating a multinomial probit model of brand choice using the method of simulated moments. *Marketing Sci.* **11**(4) 386–407.

Chintagunta, Pradeep K. 1993. Investigating purchase incidence, brand choice and purchase quantity decisions of households. *Marketing Sci.* **12** 194–208.

Chintagunta, Pradeep K. 1998. Inertia and variety seeking in a model of brand-purchase timing. *Marketing Sci.* **17**(3) 253–270.

Chintagunta, Pradeep K. 2001. Endogeneity and heterogeneity in a probit demand model: Estimation using aggregate data. *Marketing Sci.* **20**(4) 442–456.

Dillon, Willian R., Sunil Gupta. 1996. A segment-level model of category volume and brand choice. *Marketing Sci.* **15**(1) 38–59.

Dolan, R. J. 1987. Quantity discounts: Managerial issues and research opportunities. *Marketing Sci.* **6**(1) 1–23.

Erdem, Tulin. 1996. A dynamic analysis of market structure based on panel data. *Marketing Sci.* **15**(4) 359–378.

Fader, Peter S., Bruce G. S. Hardie. 1996. Modeling consumer choice among SKUs. *J. Marketing Res.* **33** 442–452.

Gelfand, A. E., A. F. M. Smith. 1990. Sampling-based approaches to calculating marginal densities. *J. Amer. Statist. Assoc.* **85** 398–409.

Guadagni, Peter M., John D. C. Little. 1983. A logit model of brand choice calibrated on scanner data. *Marketing Sci.* **2**(3) 203–238.

Hanemann, M. W. 1984. Discrete/continuous models of consumer demand. *Econometrica* **52** 541–561.

Jain, Dipak C., Naufel J. Vilcassim. 1991. Investigating household purchase timing decisions: A conditional hazard function approach. *Marketing Sci.* **10**(1) 1–23.

Kadiyali, Vrianda, Pradeep Chintagunta, Naufel Vilcassim. 2000. Manufacturer-retailer channel interactions and implications for channel power: An empirical investigation of pricing in a local market. *Marketing Sci.* **19**(2) 127–148.

Kalyanam, Kirthi, Daniel S. Putler. 1997. Incorporating demographic variables in brand choice models: An indivisible alternatives framework. *Marketing Sci.* **16**(2) 166–181.

Kamakura, Wagner, Byung-Do Kim, Jonathan Lee. 1996. Modeling preference and structural heterogeneity in consumer choice. *Marketing Sci.* **15**(2) 152–172.

Kim, Jaehwan, Greg M. Allenby, Peter E. Rossi. 2002. Modeling consumer demand for variety. *Marketing Sci.* **21**(3), 229–250.

Manchanda, Puneet, Asim Ansari, Sunil Gupta. 1999. The "shopping basket": A model for multicategory purchase incidence decisions. *Marketing Sci.* **18**(2) 95–114.

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

Montgomery, Alan L. 1997. Crating micro-marketing pricing strategies using supermarket scanner data. *Marketing Sci.* **16**(4) 315–337.

Neslin, Scott A., Caroline Henderson, John Quelch. 1985. Consumer promotions and the acceleration of product purchases. *Marketing Sci.* **4**(2) 147–165.

Newton, M. A., A. E. Raftery. 1994. Approximating Bayesian inference with the weighted likelihood bootstrap. *J. Roy. Statist. Soc. (B)* **56** 3–48.

Papatla, Purushottam. 1996. A multiplicative fixed-effects model of consumer choice. *Marketing Sci.* **15**(3) 243–261.

Pedrick, James H., Fred S. Zufryden. 1991. Evaluating the impact of advertising media plans: A model of consumer purchase dynamics using single-source data. *Marketing Sci.* **10**(2) 111–130.

Rossi, Peter E., Robert E. McCulloch, Greg M. Allenby. 1996. The value of purchase history data in target marketing. *Marketing Sci.* **15**(4) 321–340.

Silva-Risso, Jorge M., Randolph E. Bucklin, Donald G. Morrison. 1999. A decision support system for planning manufacturers' sales promotion calendars. *Marketing Sci.* **18** 274–300.

Sydsaeter, Knut, Arne Strom, Peter Berck. 1999. *Economists' Mathematical Manual*. Springer, New York.

Yang, Sha, Yuxin Chen, Greg M. Allenby. 2003. Bayesian analysis of simultaneous demand and supply. *Quant. Marketing Econom.* Forthcoming.