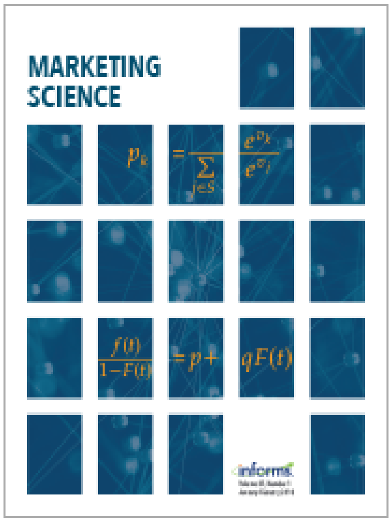


Marketing Science

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

Modeling Consumer Demand for Variety

Jaehwan Kim, Greg M. Allenby, Peter E. Rossi,

Modeling Consumer Demand for Variety

Jaehwan Kim • Greg M. Allenby • Peter E. Rossi
Leeds School of Business, University of Colorado at Boulder, Boulder, Colorado 80309
Fisher College of Business, Ohio State University, Columbus, Ohio 43210
Graduate School of Business, University of Chicago, Chicago, Illinois 60637
jaehwan.kim@colorado.edu • allenby.1@osu.edu • peter.rossi@gsb.uchicago.edu

Abstract

Consumers are often observed to purchase more than one variety of a product on a given shopping trip. The simultaneous demand for varieties is observed not only for packaged goods such as yogurt or soft drinks, but in many other product categories such as movies, music compact disks, and apparel. Multinomial (MN) choice models cannot be applied to data exhibiting the simultaneous choice of more than one variety. The random utility interpretation of either the MN logit or probit model uses a linear utility specification that cannot accommodate interior solutions with more than one variety (alternative) chosen.

To analyze data with multiple varieties chosen requires a nonstandard utility specification. Standard demand models in the economics literature exhibit only interior solutions. We propose a demand model based on a translated additive utility structure. The model nests the linear utility structure, while allowing for the possibility of a mixture of corner and interior solutions where more than one but not all varieties are selected. We use a random utility specification in which the unobservable portion of marginal utility follows a log-normal distribution. The distribution of quantity demanded (the basis of the likelihood function) is derived from these log-normal random utility errors. The likelihood function for this class of models with mixtures of corner and interior solutions is a mixed distribution with both a continuous density portion and probability mass points for the corners. The probability mass points must be calculated by integrals of the log-normal errors over rectangular regions. We evaluate these high-dimensional integrals using the GHK approximation. We employ a Bayesian hierarchical model, allowing household-specific utility parameters.

Our utility specification related to the approach of Wales and Woodland (1983) who employ a translated quadratic utility function. Wales and Woodland were only able to study, at the most, three varieties because there was no practical way to evaluate the utility function at that time. In addition, the quadratic utility specification is not a globally valid utility function, making welfare computations and policy experiments questionable. Hendel (1999) and Dube (1999) present an alternative approach in the utility function which is constructed by summing up over unobservable consumption occasions. While only one variety is consumed on each occasion, the marginal utilities of varieties change over the consumption occasions, giving rise to a simultaneous purchase of multiple varieties.

Our Bayesian inference approach allows us to obtain individual household estimates of utility parameters. Household utility estimates are used to compute the value of each variety. We compute a compensating value for the removal of each flavor; that is, we compute the monetary equivalent of the household's loss in utility from removal of a flavor. These calculations show that households highly value popular flavors and would incur substantial utility losses from removal of these flavors from the yogurt assortment.

Next we consider the implications of our model for retailer assortment and pricing policies. Given limited shelf space, only a subset of the possible varieties can be displayed for purchase at any one time. If consumers value variety, then a retailer with lower variety must compensate the consumers in some way, such as a lower price level. We see this trade-off between price and variety across different retailing formats. Discount or warehouse format retailers often have both lower variety and lower prices. To measure this trade-off, we explore the utility loss from reduction in variety and find the reductions in price that will compensate for this utility loss. These price reduction calculations must be based on a valid utility structure.

Heterogeneity in tastes is critical in these utility computations and policy experiments. We find that a relatively small fraction of households with extreme preferences dominate the compensating value computations. That is, some households are observed to purchase mostly or exclusively one variety. These households must be heavily compensated for the removal of this variety from the assortment. In some retailing contexts, customization of the assortment is possible at the customer level. We show that such customization virtually eliminates any utility loss from reduction in variety.

(*Variety; Corner Solutions; Assortment; Translated Utility*)

0732-2399/02/2103/0229/$05.00
1526-548X electronic ISSN

1. Introduction

Consumer demand for products where a variety of flavors, styles, and colors are available often results in purchases of multiple varieties. For example, households are observed to purchase more than one variety of products such as produce, meats, yogurt, and soft drinks during the same trip to the store. Outside the realm of packaged good products, demand for variety is common, e.g., entertainment products such as movies, books, and apparel. Important questions such as the selection of the optimal assortment of varieties and pricing require a rigorous utility-based model. Standard choice models are based on a linear utility structure in which only one variety will be selected at each purchase occasion. This is clearly at variance with the data. In the applied demand literature (cf. Deaton and Muelbauer 1980), demand models exhibit only interior solutions in which all varieties are purchased. In this paper, we develop a new utility-based model and estimation procedure that can accommodate a mixture of corner and interior solutions.

To allow for simultaneous demand for multiple varieties, each variety should properly be thought of as an imperfect substitute for the other varieties. Our demand model is based on a translated, nonlinear, but additive utility structure. This model provides a parsimonious specification that allows both interior and corner solutions as well as diminishing marginal utility, while nesting the standard linear utility structure. The likelihood function for this model is derived from normal random errors in marginal utility. Evaluation of the likelihood involves high-dimensional integrals of normal distributions over rectangular regions and, for this reason, has not been used in either the economics or marketing literature. Coupled with a simulation approach to evaluate the likelihood, we use a Bayesian hierarchical model of household heterogeneity. Our Bayesian approach allows for the computation of household-level parameter estimates that facilitate utility calculations. Non-Bayesian approaches to random coefficient models (cf. Nevo 2001) cannot be used to develop household-level estimates.

We apply our model to data on purchases of varieties of yogurt. In this data set, all purchases involve corner solutions and households frequently purchase more than one variety on the same shopping trip. Because our ultimate goal is to make policy recommendations about assortment and pricing, we enlarge the model to include as "outside" good—the composite of all other purchases. This allows for substitution between varieties as well as between the yogurt category and other goods. Estimates from the hierarchical model reveal differences between varieties in base preference as well as different rates of diminishing marginal utility. Moreover, households differ greatly in their preferences for varieties with some households showing extreme preferences for particular flavors.

To better understand the value households place on particular varieties, we compute a compensating value by household for the removal of each flavor. That is, we compute the monetary equivalent of the household's loss in utility from removal of a flavor. These calculations show that households highly value popular flavors and would incur substantial utility losses from removal of these flavors from the yogurt assortment.

Next, we consider the implications of our model for retailer assortment and pricing policies. Given limited shelf space, only a subset of the possible varieties can be displayed for purchase at any one time. If consumers value variety, then a retailer with lower variety must compensate the consumers in some way, such as a lower price level. We see this trade-off between price and variety across different retailing formats. Discount or warehouse format retailers often have both lower variety and lower prices. To measure this trade-off, we explore the utility loss from reduction in variety and find the reductions in price that will compensate for this utility loss. These price reduction calculations must be based on a valid utility structure.

The remainder of the paper is organized as follows. In §2, we introduce an economic model of imperfect substitutes that allows for interior solutions and illustrate its properties. Alternative explanations and models for demand for variety are discussed in

§3. In §4, we describe scanner panel data of yogurt purchases to which we apply our model. We consider both the "inside" good (allocation of expenditure on yogurt) and the "outside" good case (allocation of expenditure to yogurt category as well as among flavors) in §5. In §6, we consider the problem of valuing the variety in the assortment of yogurts via various utility-based calculations. We also consider methods of exploiting household heterogeneity to offer customized assortments. Concluding remarks are offered in §7.

2. A Model for Variety

In this section, we propose a utility function and a statistical specification for a model of horizontal variety that can accommodate both interior and corner solutions. Our approach is to retain the additive utility structure popular in the choice literature but to introduce nonlinearities and translation to achieve diminishing returns and corner solutions. It is important for the utility function to have diminishing marginal returns to capture satiation as we model demand situations where more than one unit is purchased and consumed. We retain an additive utility structure because the offerings under consideration are not jointly consumed and, hence, the utility gained from one variety is unaffected by the consumption of others. For example, there is no interaction in utility from consuming two flavors together. This makes sense for offerings within a product category that are close substitutes. However, additivity may not be appropriate for models of brands in different product categories.

2.1. Utility Structure

We define utility over the $j = 1, \ldots, m$ varieties

$$\bar{U}(x) = \sum_j \psi_j (x_j + \gamma_j)^{\alpha_j}, \tag{1}$$

where x is the vector of quantity demanded with element x_j, and ψ_j, γ_j, and α_j are parameters of the utility function. The utility in Equation (1) is an additive but nonlinear utility function. Equation (1) defines a valid utility function under the restrictions that $\psi_j > 0$ and $0 < \alpha_j \leq 1$. ψ is restricted to be positive by reparameterizing in terms of $\psi^* = \ln(\psi)$ and α is restricted to be in the unit interval by defining $\alpha^* = \ln(\alpha/1 - \alpha)$.

The utility in Equation (1) is a family of translated utility functions where γ_j controls the translation and α_j influences the rate of diminishing marginal returns. If γ_j is 0, then there will be no corner solutions as the indifference curves are tangent to the axes. However, for positive γ, the indifference curves will have finite nonzero slope at the axes, creating the possibility of a corner solution.

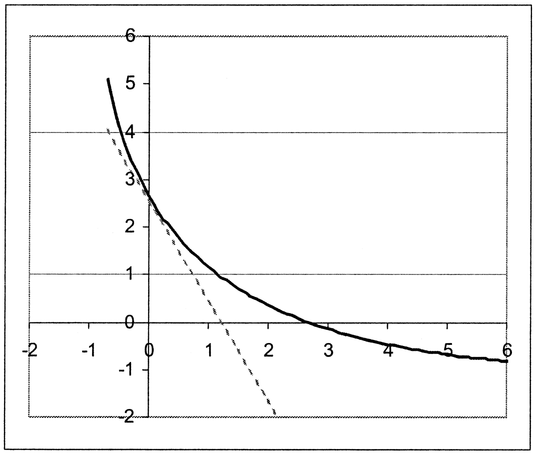
Figure 1 illustrates the two good cases. The solid lines in the figure are indifference curves. The dashed line is the budget constraint $p'x = E$, where p is the vector of prices and E is the expenditure allocation for the group of goods under study. Panel A of Figure 1 shows the situation in which there can be corner solutions ($\gamma > 0$). The left graph shows a corner solution with only good 2 consumed. In Panel B of Figure 1, $\gamma = 0$, so that only interior solutions are allowed. As illustrated, the utility function generates a flexible family of indifference curves that are capable of both interior and corner solutions.

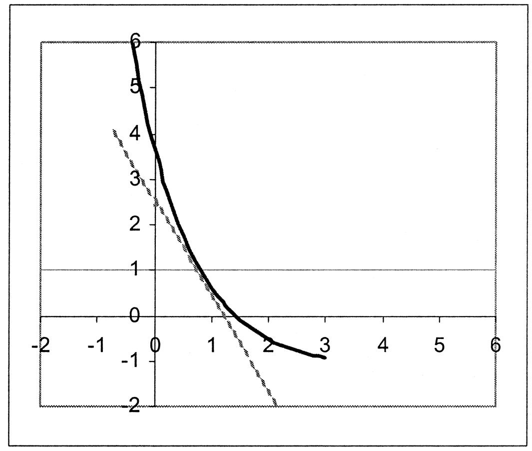
The utility specification in Equation (1) can accommodate a wide variety of situations, including the purchase of a large number of different varieties as well as purchases where only one variety is selected. If a particular variety has a high value of ψ_j and a value of α_j close to one, then we would expect to see purchases of large quantities of only one variety (high-baseline preference and low satiation). On the other hand, small values of α imply a high-satiation rate; we expect to see multiple varieties purchased if the ψ's are not too different.

For the case where the goods have well-defined characteristics, an approach to specifying a parsimonious utility function is to specify utility over characteristics and regard the goods as bundles of these characteristics (cf. Berry et al. 1995). In many important marketing situations, it is difficult, if not impossible, to define characteristics of products. Moreover, in the case of flavors, colors, and styles, these qualities are difficult to capture on a continuous scale

Figure 1 Interior and Corner Solutions with Proposed Utility Function

Panel A: $\gamma_1 = \gamma_2 = 1$

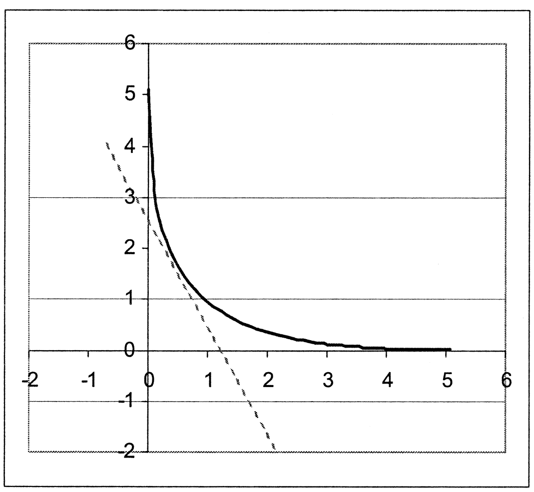


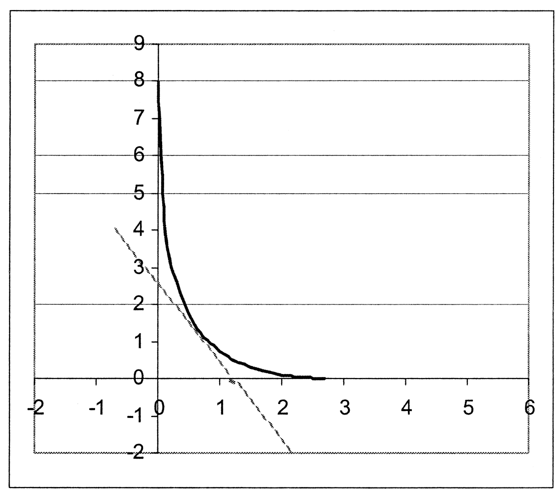


(a)
$\psi_1 = \psi_2$ and $\alpha_1 = \alpha_2 = 0.4$

(b)
$\psi_1 > \psi_2$ and $\alpha_1 = \alpha_2 = 0.4$

Panel B: $\gamma_1 = \gamma_2 = 0$





(a)
$\psi_1 = \psi_2$ and $\alpha_1 = \alpha_2 = 0.4$

(b)
$\psi_1 > \psi_2$ and $\alpha_1 = \alpha_2 = 0.4$

Note: In all cases, $p_1 = 2p_2$.

and are best thought of as distinct categories. Therefore, we adopt a specification in which each good is treated as a distinct variety.

2.2. Random Utility Specification

To develop our statistical specification, we follow a standard random utility approach and introduce a multiplicative normal error into marginal utility:

$$\ln(U_j) = \ln(\bar{U}_j) + \varepsilon_j, \qquad \varepsilon \sim N(0, 1), \tag{2}$$

where $\overline{U}_j$ is the derivative of the utility function in Equation (1) with respect to x_j. We use a log-normal error term to enforce positivity of marginal utility. We assume that ε has an identity covariance structure for simplicity. In our empirical analysis, we consider demand for different yogurt flavors, and do not feel that this assumption is too restrictive. However, for other variety problems this assumption may not be valid, and the model could be extended to include a full covariance matrix for ε.

Given the utility function in Equation (1), we derive the demand system for the set of goods under study conditional on the expenditure allocation to this set of goods. In the random utility approach, it is assumed that the consumer knows the value of ε, and that this represents omitted factors that influence marginal utility but are not observable to the data analyst. If we derive the optimal demand by maximizing utility subject to the budget constraint and conditional on the random utility error, we define a mapping from p, E, and ε to demand. Assuming a distribution for ε provides a basis for deriving the distribution of optimal demand, denoted x^*. There are two technical issues in deriving the distribution of demand: (1) Optimal demand is a nonlinear function of ε and requires use of change-of-variable calculus, and (2) the possibility of corner solutions means that there are point masses in the distribution of demand and, thus, the distribution of demand will be a mixed discrete-continuous distribution. Computing the size of these point masses involves integrating the normal distribution of ε over rectangular regions of R^m.

To solve for optimal demand, we form the Lagrangian for the problem and derive the standard Kuhn-Tucker first-order conditions. It is important to remember that the utility function specified in Equation (1), $\overline{U}$, is only the deterministic part of utility (that observed by us) and that the consumer maximizes U that includes the realization of the random utility errors. The Lagrangian is given by

$$U(x) - \lambda(p'x - E).$$

Differentiating the Lagrangian gives the standard Kuhn-Tucker first-order conditions:

$$\bar{U}_j e^{\varepsilon_j} - \lambda p_j = 0 \qquad \text{if } x_j^* > 0,$$
$$\bar{U}_j e^{\varepsilon_j} - \lambda p_j < 0 \qquad \text{if } x_j^* = 0.$$

x^* is the vector of optimal demands for each of the m goods under consideration. Dividing by price and taking logs, the Kuhn-Tucker conditions can be rewritten as:

$$V_j(x_j^* \mid p) + \varepsilon_j = \ln\lambda \quad \text{if } x_j^* > 0, \tag{3}$$
$$V_j(x_j^* \mid p) + \varepsilon_j < \ln\lambda \quad \text{if } x_j^* = 0, \tag{4}$$

where λ is the Lagrangian multiplier and

$$V_j(x_j^*|p) = \ln(\psi_j \alpha_j (x_j^* + \gamma_j)^{\alpha_j - 1}) - \ln(p_j), \quad j = 1, \ldots, m.$$

Optimal demand satisfies the Kuhn-Tucker conditions in Equations (3) and (4) as well as the "adding-up" constraint that total $p'x^* = E$. The "adding-up" constraint induces a singularity in the distribution of x^*. To handle this singularity, we use the standard device of differencing the first-order conditions with respect to one of the goods. Without loss of generality, we assume that the first good is always purchased (one of the m goods must be purchased because we assume that $E > 0$) and subtract condition 3 for good 1 from the others. This reduces the dimensionality of the system of equations by one. Equations (3) and (4) are now equivalent to

$$v_j = h_j(x^*, p) \quad \text{if } x_j^* > 0, \tag{5}$$
$$v_j < h_j(x^*, p) \quad \text{if } x_j^* = 0, \tag{6}$$

where $v_j = \varepsilon_j - \varepsilon_1$, $h_j(x^*, p) = V_1 - V_j$, and $j = 2, \ldots, m$.

The likelihood for $x^* = (x_1^*, \ldots, x_m^*)'$ can be constructed by utilizing the p.d.f. of $\upsilon = (\upsilon_2, \ldots, \upsilon_m)'$, the Kuhn-Tucker conditions in Equations (5) and (6), and the adding-up constraint $p'x^* = E$. Because ε has a multivariate normal distribution with an identity covariance matrix, $\upsilon = (\upsilon_2, \ldots, \upsilon_m)' \sim N(0, \Omega)$ where Ω is a $(m-1) \times (m-1)$ matrix with diagonal elements (i, i) equal to 2 and off-diagonal elements (i, j) equal to 1. Given that corner solutions will occur with nonzero probability, the distribution of optimal demand will have a mixed discrete-continuous distribution with lumps of probability corresponding to regions of ε that imply corner solutions. Thus, the likelihood function will have a density component corresponding to the goods with nonzero quantities and a mass function corresponding to the corners in which some of the goods will have zero optimal demand. The probability that n of the m goods are selected is equal to

$$\begin{aligned} &P(x_i^* > 0 \quad \text{and} \quad x_j^* = 0;\ i = 2, \ldots, n \quad \text{and} \\ &\quad j = n+1, \ldots m) \\ &\quad = \int_{-\infty}^{h_m} \cdots \int_{-\infty}^{h_{n+1}} \phi(h_2, \ldots, h_n, \upsilon_{n+1}, \ldots, \upsilon_m | 0, \Omega) \\ &\qquad\qquad \cdot |J| d\upsilon_{n+1} \cdots d\upsilon_m, \end{aligned} \tag{7}$$

where $\phi(\cdot)$ is normal density, $h_j = h_j(x^*, p)$, and J is the Jacobian,

$$J_{ij} = \frac{\partial h_{i+1}(x*; p)}{\partial x_{j+1}^*}, \qquad i, j = 1, \ldots, n-1.$$

The intuition behind the likelihood function in Equation (7) can be obtained from the Kuhn-Tucker conditions in Equations (5) and (6). For goods with first-order conditions governed by Equation (5), optimal demand is an implicitly defined nonlinear function of ε given by $h(\,)$. We use the change-of-variable theorem to derive the density of x^* (this generates the Jacobian term in Equation (7)). For goods not purchased, Equation (6) defines a region of possible values of ν that are consistent with this specific corner solution. The probability that these goods have zero demand, is calculated by integrating the normal distribution of ν over the appropriate region.

If there are only corner solutions with one good chosen, our model collapses to a standard choice model. The probability that only one good is chosen is given by

$$\begin{aligned} &P(x_j^* = 0,\ j = 2, \ldots, m) \\ &\quad = \int_{-\infty}^{h_m} \cdots \int_{-\infty}^{h_2} \phi(\nu_2, \ldots, \nu_m) d\nu_2 \cdots d\nu_m. \end{aligned} \tag{8}$$

Similarly, we can derive the distribution of demand for the case in which all goods are at an interior solution:

$$P(x_j^* > 0;\ i = 2, \ldots, m) = \phi(h_2, \ldots, h_m | 0, \Omega) |J|. \tag{9}$$

2.3. Heterogeneity

We introduce heterogeneity into the model by specifying a random-effects distribution for the ψ parameters

$$\psi_h^* \sim N(\bar{\psi}^*, D_{\psi^*}), \tag{10}$$

where $h = 1, \ldots, H$ indexes the households and $\psi = \exp(\psi^*)$. The multivariate normal distribution of heterogeneity for ψ_h^* is flexible as it does not restrict individual parameters to a specific support and allows elements of ψ_h^* to covary across the population. The model in Equation (10) could easily be modified to include a set of covariates in the mean function, $\bar{\psi}^* = z'\delta$, as in Rossi et al. (1996) or Ainslie and Rossi (1998). Household characteristics such as demographics might be included. These are only important to the extent that demographics matter or that the fitted model is to be used to extrapolate to a different population of households. Rossi et al. (1996) document that demographics have little explanatory power. Marketing mix variables (other than price) such as display and feature are also candidates for inclusion. Display and feature activity are the same for all varieties of the same brand, which implies that these covariates will drop out of the resulting demand functions used in our empirical analysis.

A random-effects distribution is not specified for the other model parameters because sufficient information in the data does not exist. For example, allowing households to exhibit different rates of satiation for the goods (α_j) would require long purchase histories that we do not observe. However, there are no conceptual problems specifying a random-effects distribution for all model parameters.

2.4. Inference

The model is estimated by the Markov Chain Monte Carlo method making use of the Metropolis-Hastings (M-H) algorithm to investigate the posterior distribution of model parameters (Metropolis et al. 1953, Chib and Greenberg 1995). A key advantage of our approach is that it is likelihood-based (as opposed to the GMM approach of Berry et al. 1995) or Dube (1999) and that we are able to obtain household-level estimates of the utility parameters. Details of the estimation procedure are provided in the Appendix.

The joint distribution of $(x_1^*, \ldots, x_m^*)'$ is of mixed discrete-continuous form in Equation (7) that cannot be evaluated directly. In evaluating the likelihood, we transform Equation (7) to the product of two factors as follows. Consistent with our earlier discussion, suppose that a consumer chose first n out of m items. By decomposing $\upsilon = (\upsilon_2, \ldots, \upsilon_m)'$ into $\nu_a = (\upsilon_2, \ldots, \upsilon_n)'$ and $\nu_b = (\upsilon_{n+1}, \ldots, \upsilon_m)'$ such that

$$\begin{bmatrix} \nu_a \\ \nu_b \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} \Omega_{aa} & \Omega_{ab} \\ \Omega_{ba} & \Omega_{bb} \end{bmatrix} \right), \tag{11}$$

ν_a and $\nu_b \mid \nu_a$ are normally distributed, then $\nu_a \sim MVN(0, \Omega_{aa})$ and $\nu_b \mid \nu_a = h_a \sim MVN(\mu, \Sigma)$ where $\mu = \Omega_{ba}\Omega_{aa}^{-1}h_a$, $\Sigma = \Omega_{bb} - \Omega_{ba}\Omega_{aa}^{-1}\Omega_{ab}$, and $h_a = (h_2, \ldots, h_n)'$. Then, Equation (7) can be rewritten as the product of two factors:

$$\begin{aligned} &P(x_i^* > 0 \text{ and } x_j^* = 0;\ i = 2, \ldots, n \text{ and } j = n+1, \ldots, m) \\ &\quad = \phi_{\nu_a}(h_2, \ldots, h_n \mid 0, \Omega_{aa})|J| \\ &\qquad \cdot \int_{-\infty}^{h_m} \cdots \int_{-\infty}^{h_{n+1}} \phi_{\nu_b|\nu_a}(\nu_{n+1}, \ldots, \nu_m \mid \mu, \Sigma) d\nu_{n+1} \cdots d\nu_m. \end{aligned} \tag{12}$$

We use the GHK simulator (Keane 1994, Hajivassiliou et al. 1996) to evaluate the multivariate normal integral that appears in Equation (12).

3. Alternative Explanations and Modeling Approaches

The variety-seeking literature in marketing (McAlister and Pessemier 1982, McAlister 1982, Lattin and McAlister 1985, Erdem 1996, Kahn et al. 1986) focused on temporal changes in tastes from purchase occasion to purchase occasion. This literature is aimed at explaining switching in brand choice, and employs models that assume a linear utility structure with temporal dependence. In contrast, this paper is aimed at explaining simultaneous purchase for multiple varieties, which we term "horizontal" variety seeking (see also Bradlow and Rao 2000).

We model the demand for varieties with a direct utility function that views the varieties as imperfect substitutes. That is, we view households as acting in purchase decisions as though they have a representative direct utility function defined over the varieties. We are the first to implement a feasible likelihood-based approach for these types of utility functions. Wales and Woodland (1983) use a quadratic utility function and note that likelihood evaluation was infeasible with more than a few varieties. The purchase incidence models such as those considered by Hanneman (1984), Chiang (1991), Chintagunta (1993), and Arora et al. (1998) can be viewed as a two-good special case of our situation.

An alternative to our direct approach is to view the purchases as resulting from a stream of consumption occasions. Purchase demand for multiple varieties then stems from changes in the marginal utility for each variety over these consumption occasions as in Walsh (1995). Dube (1999) (see also Hendel 1999) pursues a simplified version of this approach. In Dube's model, households have a linear utility function for each consumption occasion with varying utility weights. The linear utility for each consumption occasion is raised to a power and summed over consumption occasions. The model is completed by making random effect distributional assumptions regarding the number of consumption occasions (assumed to be Poisson) and the variation in the weights (independent normal errors).

The contribution of the Hendel/Dube approach is to derive a purchase demand system by aggregating over *unobserved* consumption occasions. There are many such models that might be postulated by changing the distribution of the number of purchase occasions, the distribution of random utility weights, or the consumption occasion utility function. Each possibility gives rise to a different purchase demand system of equations. In our view, Hendel/Dube provides an alternative to our approach, and it is entirely an empirical question as to which model better fits the data. Our approach can be viewed as more flexible than Hendel/Dube because we allow for multiple curvature or satiation parameters, one for each variety. As we demonstrate below, our data clearly call for different satiation parameters, which make both a substantive and statistical difference in the conclusions.

There is also a difference in the inference methods used by us and by Dube. We employ a likelihood-based approach while Dube uses a method of moments. This means that Dube's estimators of common parameters will be inefficient relative to a full likelihood approach. However, our likelihood approach comes at a cost. We employ the GHK simulator to evaluate the integrals required for evaluation of the likelihood. In the size panel, data sets commonly used in marketing, would not be able to handle more than 15 to 20 different varieties. Finally, we use a full Bayesian approach that uses data augmentation to obtain household estimates (see Allenby and Rossi (1999) for a complete discussion of Bayesian methods for obtaining household parameter estimates). This means that we can make utility loss calculations at both the household and aggregate level—something Dube cannot do.

Another approach to modeling direct utility would be to define the utility function for the household as stemming from some aggregation of the utility functions for each household member. The problem here is that we do not observe the consumption for each member of the household and there is little guidance as to how to aggregate the utility functions for each member into a household utility function.

4. Data

We estimate our model using scanner-panel data from the yogurt category. We examine demand for five popular flavors of Dannon yogurt (blueberry, mixed berry, piña colada, plain, and strawberry) in the popular 8-ounce size. In our data, there is virtually no brand switching between the Dannon and Yoplait yogurt brands. Less than 2% of the observations involved purchase of both brands. Therefore, we restrict attention to the dominant market share Dannon brand varieties. However, there is nothing conceptually different about applying our utility specification to multiple varieties from multiple brands.

We start with the data set of all 8-ounce Dannon yogurt purchases from all households (this data set has 7,622 observations). We then restrict attention to only those households with more than one purchase occasion for at least one of our five varieties and for whom information on total grocery expenditure is available. This yields a data set with 332 households and 2,380 purchase observations. We consider only those purchase occasions on which one of our five flavors are purchased. That is, we are not attempting to explain the incidence of purchase in the yogurt category. This would require a more elaborate model in which the marginal utility of yogurt is time varying due to dynamic considerations (such as changes in household inventory and anticipation of price discounts) that is beyond the scope of this paper. Our basic utility model, however, is one that could be used to form the basis of a rigorous dynamic structural approach.

Table 1 presents information on the frequency of corner and interior solutions. A corner solution is defined as a purchase occasion on which only one variety is purchased. Table 1 clearly shows the problem with application of standard choice models because more than 20% of the purchase occasions involve the simultaneous purchase of more than one variety. We can eliminate the possibility that household "portfolio" effects are the only possible explanation for the simultaneous purchase of multiple varieties. The 69 single-person households in our

Table 1 Frequency of Corner and Interior Solutions in the Yogurt Data

Flavor	Purchase Incidence	Corner Solution	Interior Solution
Strawberry	989	571	418
Blueberry	545	309	236
Piña Colada	570	281	289
Plain	361	338	23
Mixed Berry	570	325	245
Total	2380	1824 (76.6%)	556 (23.4%)

data set have 464 purchase occasions with the same frequency of interior solutions (26%).

Tables 2 and 3 provide information on the quantity of yogurt purchased. Table 2 reports the distribution of purchase quantity. More than 50% of the purchase occasions involve quantities of at least 2 units, indicating that the data is capable of providing information about the flavor satiation. Table 3 displays the incidence of corner and interior solutions for purchase occasions in which at least two units are demanded. This table shows that plain and mixed berry yogurt are more often purchased in isolation, while the other flavors are typically purchased in conjunction with the other flavors. One possible explanation for the differing behavior of plain and mixed berry varieties is that they are strongly preferred by some households. Alternatively, consumers may not tire of these varieties as rapidly as the others. The model discussed above has both features and will allow us to distinguish between these effects.

5. Calibration of Inside and Outside Good Models

Our ultimate goal is the measurement of the utility that households derive from having a variety of goods from which to choose. The extent to which any given flavor is valued depends on the degree to which this flavor is substitutable for other flavors and other goods. For example, if a given flavor becomes relatively more expensive or is no longer available, then we expect substitution to other yogurts and possibly other grocery items. We will start the investigation of substitution patterns by restricting attention to only yogurt flavors. We will then enlarge the model to include a composite "outside" good consisting of expenditures on all other items in the grocery store.

Table 2 Distribution of Purchase Quantity

Purchase Quantity	Frequency	%
1	1140	47.90
2	745	31.30
3	228	9.58
4	133	5.59
5	42	1.76
6	60	2.52
7	8	0.34
8	17	0.71
9	1	0.04
10	3	0.13
11	1	0.04
12	1	0.04
21	1	0.04

5.1. Inside Good Model

We first consider an inside good model in which demand is allocated among the five varieties of yogurt, conditional on the total expenditure on the yogurt category. The likelihood is derived by assuming that on each purchase occasion the households maximizes utility of the form $U(x) = \Sigma_{j=1}^{5} \psi_j (x_j + \gamma_j)^{\alpha_j}$ subject to the constraint that $\Sigma_{j=1}^{5} p_{j,ht} x_j \leq M_{ht}$, where M is the expenditure on yogurt. The translation parameters, γ_j, serve to allow the indifference curves to in-

Table 3 Frequency of Corner and Interior Solutions for Multiple Unit Purchases

Flavor	Observations	Corner Solution	Interior Solution
Strawberry	336	208 (0.62)	128 (0.38)
Blueberry	207	129 (0.62)	78 (0.38)
Piña Colada	183	99 (0.54)	84 (0.46)
Plain	117	108 (0.92)	9 (0.08)
Mixed Berry	198	140 (0.71)	58 (0.29)

Table 4 Parameter Estimates—Inside Goods Only (Posterior Standard Deviations)

	Common Parameters	
Flavor	$\bar{\psi}^*$	α^*
Strawberry	0.57 (0.01)	0.32 (0.06)
Blueberry	−0.15 (0.07)	0.66 (0.12)
Piña Colada	0.12 (0.09)	0.02 (0.13)
Plain	−1.54 (0.12)	3.62 (0.50)
Mixed Berry	0.00**	0.74 (0.11)

	Covariance/*Correlation* Matrix (ψ_h^*)			
Flavor	Strawberry	Blueberry	Piña Colada	Plain
Strawberry	0.14 (0.01)	*−0.08*	*−0.07*	*−0.14*
Blueberry	−0.03 (0.02)	0.86 (0.12)	*−0.18*	*0.21*
Piña Colada	−0.03 (0.02)	−0.16 (0.08)	0.92 (0.11)	*0.25*
Plain	−0.11 (0.05)	0.34 (0.13)	0.47 (0.14)	3.90 (0.38)

Note. $\psi = \exp(\psi^*)$, $\alpha = \exp(\alpha^*)/(1+\exp(\alpha^*))$ in Equation (1).
**Fixed for identification.

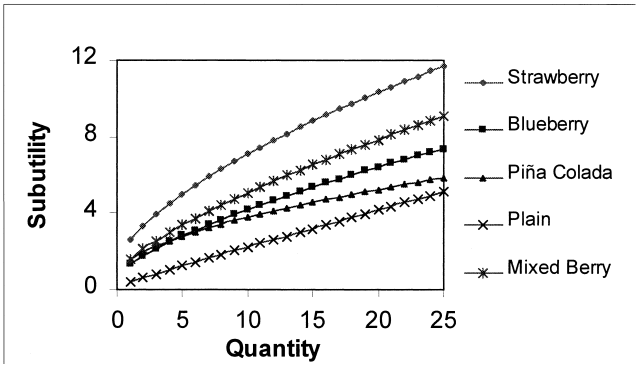
tersect the axes. Both γ and α govern the slope of the indifference curves at the point of intersection with the axes and are difficult to identify separately. For this reason, we fix $\gamma_j = 1.0$ for all j.

Table 4 presents estimates of the reparameterized α, ψ parameters (all posterior quantities are estimated with the last 5,000 draws from a sequence of 20,000 draws). Given the low autocorrelation in the MCMC chain, 5,000 draws provide accurate estimates of the posterior moments reported in Table 4. We can interpret α as influencing the rate at which households become satiated with the flavor. Larger values of the reparameterized α indicate models closer to a linear utility specification that exhibits no satiation or "wearout." Our estimates indicate large differences among varieties in the degree of "wearout." Plain yogurt has much less satiation than the other flavors.

The model of Hendel (1995) has a common curvature parameter for all goods in the utility function. Our data set illustrates that this can be restrictive. We estimated a version of our model with the restriction that all of the satiation parameters are equal across brands. This model fits the data less with a marginal log-likelihood value of −4317.62 versus an unconstrained value of −4262.28. We note that this implies Bayes Factor or Odds Ratio values overwhelmingly against the restriction. Not only is the restriction of a common curvature model rejected on statistical grounds, we also find large substantive differences between restricted and unrestricted models. Price elasticities differ markedly between the restricted and unrestricted models. In the restricted model, plain yogurt is not allowed to exhibit a lower satiation rate. This would imply that the price elasticity for plain yogurt should be greatly affected by the imposition of the restriction on the curvature parameters. The price elasticity for plain yogurt in the unrestricted model is −1.88, while in the restricted model the elasticity is −1.23, a difference of more than 50%. Moreover, the role of the curvature parameter in our highly nonlinear model is complex. In summary, differences of more than 25% are observed for 4 out of the 5 possible varieties.

The ψ_j^* parameters influence the base level of utility for flavor j. They play a role analogous to brand intercepts in the standard linear utility models. Figure 2 plots the subutility functions for each of the flavors, using the mean level of the ψ^* parameter. Our estimates clearly show nonlinearities in the utility function, rejecting the use of common linear utility functions as in Berry et al. (1995) or Besanko et al. (1998).

Figure 2 Subutility Functions—Inside Good Model



We recognize that different households have different baseline levels of utility for various flavors. The mean level of the distribution of ψ_h^*, reported in the top panel of Table 4, shows that the average household does not like plain yogurt. Of course, this is somewhat misleading as there is considerable heterogeneity in preferences for flavors. The bottom panel of Table 4 reports distribution of the ψ_h^* parameters. Plain has the highest variance across households, indicating that there are some households with extreme preferences vis-à-vis plain yogurt. Even the blueberry and piña colada flavors exhibit a good deal of heterogeneity. Heterogeneity in household preferences plays a key role in the evaluation of the policy experiments considered in §5 below. The correlation matrix reported shows small intercorrelations indicating that a characteristics-style model might not be useful for this set of flavors.

5.2. Outside Good Model

Recognizing that there is substitution not only among yogurt flavors but to other grocery products, we enlarge our model to include a composite good that consists of expenditures on all other grocery products. In our utility framework, it is natural to introduce the outside good as a sixth product with a translation parameter of 0 (insuring that there is always an interior solution with respect to this good):

$$\bar{U}(x) = \sum_{j=1}^{5} \psi_j (x_j + 1)^{\alpha_j} + \psi_{out} x_{out}^{\alpha_{out}}. \tag{13}$$

We can enlarge our data to include expenditures on other products and consider the price of the composite outside good to be the numeraire (set to \$1.0). Estimation of the unrestricted version of the above model proved to be demanding of the data. Because the basic feature of the data is that households purchase many more units of the outside good than yogurt, we are forced into a region of the parameter space in which the satiation is high for the yogurt flavors and low for the outside good. It should also be noted that the rate of satiation (second derivative of the utility function) is influenced by both α and ψ:

$$U''(x) = \psi\alpha(\alpha - 1)(x + 1)^{\alpha - 2}.$$

For these reasons, we found it difficult to estimate separate α parameters for each of the flavors in the model with the outside good. This does not mean that we cannot identify different satiation rates, but that flexibility induced by flavor-specific α's is not needed. For this reason, we restrict each of the flavor α's to be the same, while allowing for a different outside good α.

Table 5 presents the parameter estimates for the model with the outside good. As expected, the rate of satiation for the yogurts is dramatically higher than the outside good. It is also clear that the utility for the outside good is far from linear. We see a similar pattern of preferences between flavors and a similar pattern of heterogeneity in which plain yogurt preferences are most variable across households. The correlation matrix of the ψ parameters has the interesting feature that there are now some positive correlations between the outside good baseline parameter and the yogurt flavors. This can be interpreted as revealing that those households who purchase a lot of the outside good tend to dislike mixed berry relative to the other flavors. As mixed berry could be interpreted as a "risk averse" portfolio

Table 5 Parameter Estimates- Outside Good Model (Posterior Standard Deviations)

Flavor	Common Parameters $\bar{\psi}^*$	Common Parameters α^*
Strawberry	0.49 (0.02)	
Blueberry	−0.24 (0.08)	
Piña Colada	−0.15 (0.08)	−3.34 (0.07)
Plain	−1.32 (0.16)	
Mixed Berry	0.00**	
Outside Good	0.12 (0.11)	0.36 (0.08)

Flavor	Covariance/*Correlation* Matrix (ψ_h^*) Strawberry	Blueberry	Piña Colada	Plain	Outside Good
Strawberry	0.37 (0.03)	*0.25*	*0.07*	*−0.01*	*0.27*
Blueberry	0.20 (0.05)	1.77 (0.17)	*0.07*	*0.17*	*0.43*
Piña Colada	0.05 (0.05)	0.10 (0.10)	1.18 (0.13)	*0.30*	*0.35*
Plain	−0.01 (0.10)	0.52 (0.24)	0.74 (0.18)	5.14 (0.50)	*0.59*
Outside Good	0.12 (0.03)	0.40 (0.08)	0.27 (0.06)	0.96 (0.14)	0.50 (0.06)

Note. $\psi = \exp(\psi^*)$, $\alpha = \exp(\alpha^*)/(1 + \exp(\alpha^*))$ in Equation (1).
**Fixed for identification.

Table 6 Derivatives of Utility—Outside Good Model (Standard Deviation)

Flavor	1st derivative	2nd derivative
Strawberry	0.044 (0.029)	−0.033 (0.028)
Blueberry	0.033 (0.041)	−0.025 (0.032)
Piña Colada	0.035 (0.038)	−0.028 (0.030)
Plain	0.029 (0.066)	−0.017 (0.041)
Mixed Berry	0.030 (0.008)	−0.027 (0.011)
Outside Good	0.106 (0.082)	−0.004 (0.011)

choice, it is possible that households with larger shopping baskets can afford the luxury of more variety in yogurt flavors.

The fact that restricting the α parameters for the yogurts to be the same still allows for flexibility in satiation is demonstrated in Table 6. Table 6 provides the average value of the first and second derivatives of the utility function, averaged over all purchase occasions in the data. We see much smaller second derivatives for the outside good than the yogurt varieties. We still see substantial differences in the second derivatives with plain yogurt exhibiting a second derivative half the size of strawberry.

Table 7 presents price elasticities estimates for the inside only and outside good models. These are price elasticities for aggregate demand formed by integrating over the distribution of household parameter estimates. While the parameters differ substantially in magnitude between the inside and outside good models (compare Tables 4 and 5), the implied price elasticities are roughly equivalent. The addition of the outside good opens up greater substitution possibilities and should result in larger price elasticities as is clearly shown in Table 7. The availability of the outside good also affects the pattern of brand competition. In the inside good model, we force the consumer to substitute another yogurt flavor. With the addition of the outside good, an increase in the

Table 7 Own Price Elasticities for Inside/Outside Good Models

Flavor	Inside Good Model	Outside Good Model
Strawberry	−2.21	−3.38
Blueberry	−2.26	−2.71
Piña Colada	−2.56	−2.85
Plain	−1.88	−1.98
Mixed Berry	−3.53	−3.38

price of yogurt will induce substantial substitution to the outside good. Consumers will do little brand switching in the presence of the outside good, a phenomenon noted by Chintagunta (1993) and others.

We will now take our model that allows for substitution between flavors and to the outside good and examine the problem of valuation of the variety of this assortment of flavors.

Table 8 Impact on the Demand of Outside-Good

Flavor Deleted from Assortment	Sales of Outside Good	Change (%)
Strawberry	92,347	1.0
Blueberry	91,932	0.5
Piña Colada	91,930	0.5
Plain	91,780	0.3
Mixed Berry	91,936	0.5

6. Policy Implications

Demand systems derived from valid utility functions can be used to make policy recommendations in pricing and assortment. Without a valid utility structure, we must rely on reduced form models in which the demand for yogurt is linked in an ad hoc manner to the assortment of yogurts available. Even if we could agree on the nature of the link between yogurt expenditure and variety as well as a measure of variety, estimation of such a reduced form would require time series variation in variety offered that is not present in our data. Note that share or choice models have no implications for the effect of variety on total yogurt expenditure. A utility-based model allows us to analyze counterfactual experiments such as what would demand be in the absence of a particular variety. In this section, we explore the policy experiment of deleting a flavor from the assortment and use direct utility calculations to determine the utility loss and compensating value.

All retailers and manufacturers face the problem of selecting the optimal assortment of varieties to manage in a category. In addition, retailers face the general strategic problem of trading off variety for some other attribute such as price. Some store formats have high variety and relatively higher prices while others have lower variety and lower prices (e.g., warehouse clubs).

6.1. Compensating Value

The value of an assortment can be determined by computing the compensating value of adding or removing an offering from the product line. As the assortment is altered, the utility attainable for a fixed level of expenditure changes. The compensating value is the amount that the budgetary allotment would have to increase or decrease to yield the same level of utility as that attained prior to any change in the assortment. As utility is measured on an arbitrary scale, we convert to the monetary scale of compensating value for interpretability purposes.

In this section, we undertake compensating value calculations by deleting flavors from the assortment. The removal of a flavor from a product line will result in a decrease in the attainable utility of all consumers, and is affected by factors such as whether other offerings are considered good substitutes and the marginal utility of consumption. These factors are reflected in more than one of the model parameters in Equation (1), and depend on the current level of expenditure. The compensating value depends on the set of substitutes available. Deletion of a flavor will induce substitution to other flavors as well as to the outside good. For this reason, we use the outside good model from §4 in our calculations.

Table 8 provides some evidence on the shift in demand that will occur if a flavor is deleted. This demand is calculated by maximizing utility at the household level with flavors deleted and adding this demand up to the aggregate level. Table 8 shows a relatively large degree of substitution to the outside good when flavors are deleted. Thus, it is important to include the outside option in policy calculations.

To compute the compensating value, we must numerically evaluate the indirect utility function and compute the increase in expenditures required to attain the level of utility derived from the full assortment. Because prices and total grocery expenditure vary across purchase occasions, we must undertake these computations observation by observation and then sum them for each household.

Define the indirect utility function for each observation:

$$V_{ht}(p_{ht}, E_{ht}) = \max_x U(x \mid \psi_h, \bar{\alpha})$$
$$\text{s.t. } p'_{ht}x = E_{ht}.$$

U is defined over all five flavors and the outside good as in Equation (13), which means that p and x include six elements. E_{ht} is the total grocery expenditure for household h at purchase occasion t. We condition on the posterior means of the utility function parameters denoted $\overline{\psi}_h$, $\overline{\alpha}$.

To find the compensating value for deletion of flavor i, we solve for Compensating Value (CV) in

$$V_{ht}(p_{ht}, E_{ht}) = V_{ht}^{(i)}(p_{ht}, E_{ht} + CV_{ht}^{(\mathrm{i})}),$$

where $V_{ht}^{(i)}$ is defined by

$$V_{ht}^{(\mathrm{i})}(p_{ht}, E_{ht}) = \max_x U(x|\bar{\psi}_h, \bar{\alpha})$$
$$\text{s.t. } p'_{ht}x = E_{ht} \quad \text{and} \quad x_i = 0.$$

The indirect utility functions can be evaluated by numerical optimization methods (NAG routine E04UCF was used with analytic derivatives). These utility calculations are made conditional on estimates of household-level parameters that are not available using the approaches of BLP or Dube (1999).

$CV_{ht}^{(i)}$ is defined as amount by which expenditure must be increased to compensate household h on purchase occasion t so that utility will remain unaffected by the deletion of a flavor. If a flavor has unique value with poor substitutes, then the compensating value will be high. In addition, some households may have an extreme preference for a given flavor that may also cause the CV to be large. We sum up the $CV_{ht}^{(i)}$ over purchase occasions to the household level, $CV_h^{(i)} = \Sigma_{t=1}^{T_h} CV_{ht}^{(i)}$. To provide a basis for household comparison and to provide a more meaningful metric, we divide the household level

Table 9a Compensating Value Calculations

(i) CV

Flavor Deleted from Assortment	Mean	Standard Deviation
Strawberry	1.506	5.98
Blueberry	0.908	3.58
Piña Colada	0.954	5.99
Plain	0.625	2.43
Mixed Berry	0.965	8.34

(ii) PCV

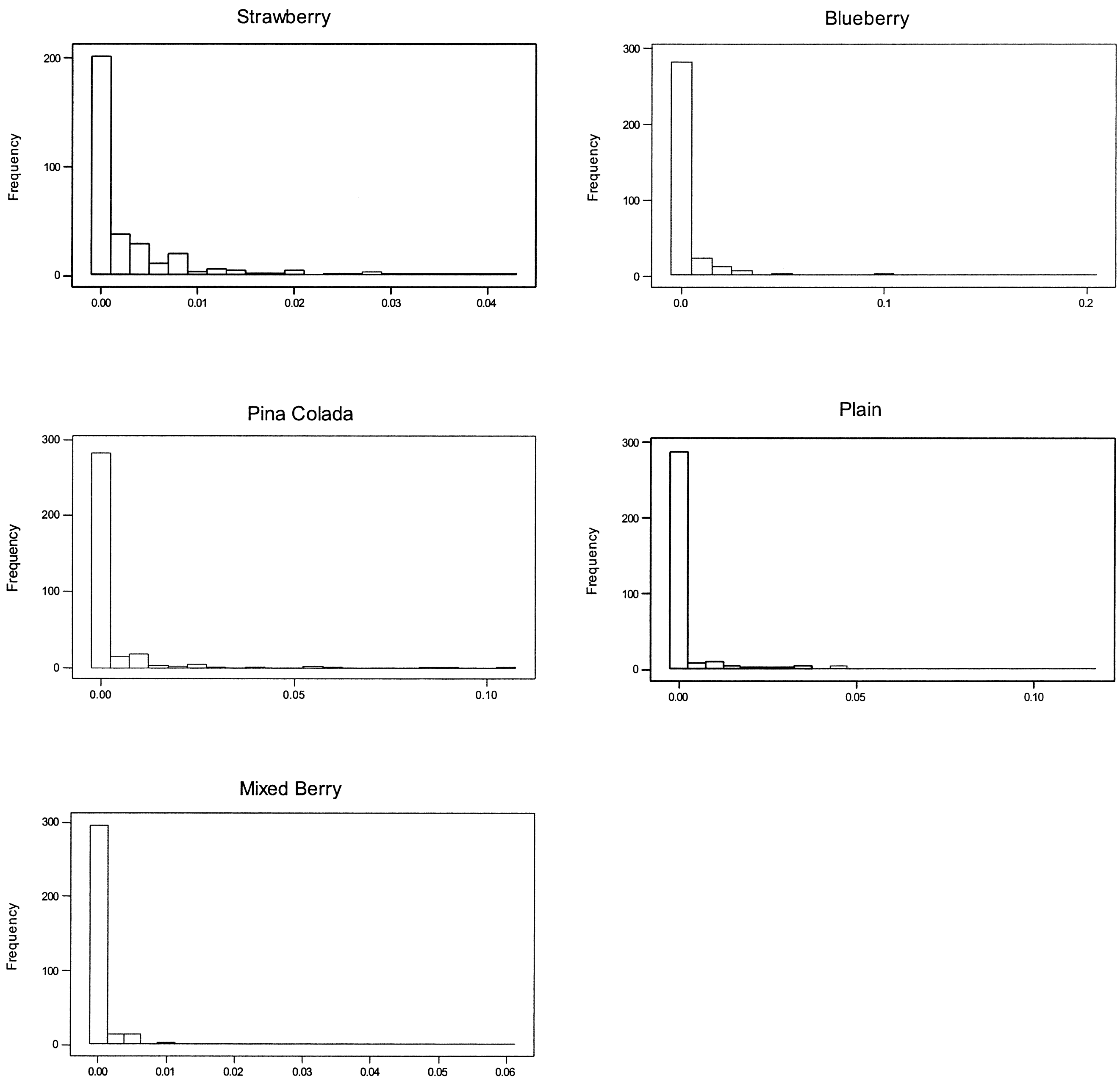
Flavor Deleted from Assortment	Mean	Standard Deviation
Strawberry	0.0034	0.007
Blueberry	0.0046	0.017
Piña Colada	0.0030	0.011
Plain	0.0033	0.012
Mixed Berry	0.0012	0.006

CV by total expenditure to express this as a percentage of expenditure:

$$PCV_h^{(i)} = \frac{CV_h^{(i)}}{\sum_{t=1}^{T_h} E_{ht}.}$$

Table 9a provides the results for the household level CV (panel *i*) as well as the percentage CV (panel *ii*). The CV calculations suggest that, on average, the utility loss from deletion of a flavor is worth about \$1 to each household over the average duration of the household in the panel (about 1 to 1.5 years). This ranges from \$1.51 for strawberry to \$.63 for plain yogurt. Our view is that these are pretty large numbers because we are deleting only one (albeit popular) SKU in the grocery store. However, the averages may be misleading because there are huge differences across households as indicated by the large standard deviations. Many households experience small utility loss while some experience large utility losses and require a much larger CV. Some households may be in the panel much longer than others. The PCV (panel *ii*) measure normalizes both for the number of purchases as well as the total volume of purchases. Figure 3 provides histograms of

Figure 3 Compensating Values as a Percent of Total Household Expenditure



PCV for each of the deleted flavors. In spite of this normalization used in the *PCV*, we still see large differences between households that are not removed by normalization.

All of the indirect utility computations used to produce our *CV* and *PCV* estimates assume that the consumer can purchase noninteger quantities of the yogurt varieties. This allows us to use standard

Table 9b Compensating Value Calculations with Integer Constraints

(i) CV		
Flavor Deleted from Assortment	Mean	Standard Deviation
Strawberry	1.285	4.68
Blueberry	0.682	2.50
Piña Colada	0.775	4.81
Plain	0.577	2.21
Mixed Berry	0.750	6.34

(ii) PCV		
Flavor Deleted from Assortment	Mean	Standard Deviation
Strawberry	0.0029	0.006
Blueberry	0.0037	0.013
Piña Colada	0.0026	0.009
Plain	0.0029	0.010
Mixed Berry	0.0010	0.005

nonlinear programming algorithms and is compatible with the above likelihood function derivation. However, one could argue that this is inappropriate in that households are constrained to purchase integer (including 0) quantities. In particular, one might expect that the integer constraint is most important in situations in which the household might wish to purchase small amounts of many varieties. In this case, the minimum purchase quantity of one 8-ounce unit might be important. On the other hand, one might argue that if, from a utility point of view, a finer grid of quantity choices was desirable, then we would see manufacturers offering smaller units. One way of assessing the magnitude of this problem is to resolve the CV problem with integer constraints. Table 9b provides the results of these calculations. The table clearly shows that imposition of integer constraints makes little difference.

The compensating value calculations are made at the household level and illustrate the important differences between households in their valuation of varieties. In most retailing contexts, customization to the household level is not possible. Retailers adopt uniform assortment and pricing policies. In the next section, we consider the changes required in a uniform pricing policy to compensate for loss in variety.

6.2. Implications for Pricing Policy

As a policy matter, retailers face space constraints and must offer a more limited variety than possible. If a retailer were to delete a flavor from a store inventory, utility losses might drive customers away from the yogurt category and the store. For this reason, we might consider a policy experiment in which the retailer compensates customers for reduced variety by lowering prices either storewide or in the yogurt category. This strategy has been adopted by warehouse format competitors such as Walmart; less variety is offered with lower prices.

To compute the price reductions necessary to compensate for loss in variety, we again use the indirect utility functions previously defined. To calculate utility-compensating price reductions, we solve for the percentage reduction in prices required to restore aggregate utility to the level present prior to its deletion from the product line. In performing this computation, we integrate over the observed purchases and posterior distribution of model parameters reported in Table 5, including the distribution of heterogeneity. We will perform two types of calculations: (1) compute the amount by which we must reduce the price of the remaining yogurts to compensate for loss of a flavor, and (2) calculate the amount by which all store prices (outside good as well as yogurt) must be reduced to compensate for loss of variety. Here we assume that overall grocery expenditure will not be affected by deletion of a yogurt flavor.

Table 10 presents results on the price reductions necessary to compensate for the loss of a flavor, using only yogurt prices. That is, the retailer lowers prices of yogurts alone to compensate for loss of variety without changing the prices of other goods in the store (as captured by the outside good). We solve the following problem to determine the level by which we must reduce yogurt prices, $r^*_{i,yogurt}$, to compensate for deletion of flavor i:

$$\text{find } r^*_{i,yogurt} \quad \text{such that } \bar{V}^{(i)}(r^*_{i,yogurt}) = \bar{V},$$

where

$$\bar{V} = \sum_{h=1}^{H}\sum_{t=1}^{T_h} V_{ht},$$

$$V_{ht} = \max U(x|\bar{\psi}_h, \bar{\alpha})\text{s.t. } x'_{yogurt}p_{yogurt,ht} + P_{out,ht}x_{out} = E_{ht},$$

and

$$\bar{V}^{(i)}(r^*_{i,yogurt}) = \sum_{h=1}^{H}\sum_{t=1}^{T_h} V^{(i)}_{ht}(r^*_{i,yogurt}),$$

$$V^{(i)}_{ht}(r^*_{i,yogurt}) = \max U(x \mid \bar{\psi}_h, \bar{\alpha}) \text{ s.t}$$

$$(x'_{yogurt}p_{yogurt,ht})r^*_{i,yogurt} + p_{out,ht}x_{out} = E_{ht} \quad \text{and} \quad x_i = 0.$$

In the top panel of Table 10, we display results that take into account the distribution of flavor preferences by integrating over the distribution of household parameters. These results show that deletion of all flavors (except for mixed berry) result in losses in utility that require substantial price reductions (15 to 21%). Plain yogurt requires the highest price cut even though, on average, households do not like plain yogurt. The reason is that there is a subset of households that highly value plain yogurt. In a traditional retailing environment, the only way the retailer can compensate for this is to substantially lower prices on the remaining yogurts.

These price reduction results all hinge on the distribution of household heterogeneity. On average, households regard the yogurt flavors as quite substitutable, but this statement is misleading. Many households have a decided preference for one flavor that drives these utility and price reduction results. To illustrate this point, we consider the price reduction experiment conditional on the average value of the baseline preference parameters, $\overline{\psi}$ in the bottom panel of Table 10. Here we see that, conditional on $\overline{\psi}$, there are only negligible reductions in utility, which means that we do not have to lower prices at all to compensate for a loss of a flavor. In models with nonlinear utility functions, any sort of welfare calculation such as the compensating value computation undertaken here is sensitive to how household heterogeneity is handled. We cannot simply "plug in" the mean value (over all households) of the relative preference vector (ψ).

Table 10 Price Reductions Required to Compensate for Utility Loss: Yogurt Prices Only

Flavor Deleted from Assortment	Loss in Utility	Compensating Value as % of Price
Conditional on Household-Level Estimates of ψ		
Strawberry	35.55	16.94
Blueberry	32.61	14.81
Piña Colada	33.54	14.96
Plain	45.26	21.25
Mixed Berry	5.25	2.25
Conditional on Mean of ψ		
Strawberry	0.2009	0.0
Blueberry	0.0002	0.0
Piña Colada	0.0007	0.0
Plain	0.0000	0.0
Mixed Berry	0.0022	0.0

Table 11 presents the results of price reduction calculations in which the whole vector of prices is reduced by the same percentage amount. Mathe-

Table 11 Price Reductions Required to Compensate for Utility Loss: All Prices

Flavor Deleted from Assortment	Loss in Utility	Compensating Value as % of Price
Strawberry	35.55	0.47
Blueberry	32.61	0.43
Piña Colada	33.54	0.44
Plain	45.26	0.60
Mixed Berry	5.25	0.07

matically, this can be expressed as find r_i^* such that $\overline{V}^{(\mathrm{i})}(r_i^*) = \overline{V}$ where

$$\bar{V} = \sum_{h=1}^{H}\sum_{t=1}^{T_h} V_{ht},$$
$$V_{ht} = \max U(x|\bar{\psi}_h, \bar{\alpha})$$
$$\text{s.t. } x'_{yogurt}p_{yogurt,ht} + p_{out,ht}x_{out} = E_{ht},$$

and

$$\bar{V}^{(i)}(r_i^*) = \sum_{h=1}^{H}\sum_{t=1}^{T_h} V_{ht}^{(i)}(r_i^*),$$
$$V_{ht}^{(i)}(r_i^*) = \max U(x|\bar{\psi}_h, \bar{\alpha}),$$
$$\text{s.t. } (x'_{yogurt}p_{yogurt,ht})r_i^* + r_i^* p_{out,ht}x_{out} = E_{ht} \quad \text{and} \quad x_i = 0.$$

The loss of yogurt flavors can be compensated for by a reduction in the overall price level in the store of around 0.5%. Our view is that this is a relatively large price reduction necessary to compensate for utility losses in one small category of the store. This illustrates the sense in which the outside good is not a good substitute for yogurt.

6.3. Customization

We have seen that fairly large price reductions are required to compensate for the loss of popular yogurt flavors. This result is driven almost exclusively by heterogeneity in preferences for yogurt flavors. Some households are made much worse off by deletion of a flavor and require rather large price reductions to keep them happy. In the traditional retailing format, in which all customers receive the same price, these results illustrate a powerful force that requires retailers to have a large variety. This variety is achieved at a high cost. We should point out that price reductions computed above are designed to insure that total utility remains the same after the deletion of a flavor. This is achieved by lowering prices to all customers by the same percentage. It can be argued that these price reductions are inefficient and possibly ineffective. By lowering prices, we make the households who prefer the remaining flavors better off and reduce but do not eliminate the utility losses of those households who prefer the deleted flavor. It is possible that those households who still experience a large utility loss will leave the yogurt category all together or, even worse, leave this retailer for another with greater variety. There are two ways to avoid this problem inherent with a uniform assortment and pricing policy: (1) adopt a customized pricing policy in which the households received customized offers or coupons to compensate for loss of variety, and (2) customize the assortment at the household level. In this section, we pursue the later suggestion of customized assortment.

If we were able to customize the assortment to each household, we should be able to reduce the utility loss incurred by reduction in the size of the assortment. In Web retailing, this is a real possibility. A Web retailer can have the full assortment of varieties, but displaying this information to the buyer may be costly in terms of navigation of the Web site and purchase decisions. For example, a Web retailer such as Netgrocer or Peapod could offer its customers the full array of 25 or more Dannon yogurt flavors. The danger here is that the customers incur a larger search and ordering cost than a customer facing a much more limited variety at the standard brick and mortar retailer. One way of avoiding these information-processing mental costs is to customize the assortment based on past purchase behavior.

To illustrate the value of customized assortments, we return to our utility loss calculations. Instead of deleting the same flavor across all households and calculating the utility loss and compensating value, we now allow the flavor deleted to vary across households. We delete the flavor that will incur the least utility loss for each household. We find that if we allow this increased flexibility, virtually no utility losses are incurred (the compensating value = 0 to 4 decimal places). This is due to the heterogeneity in households that we are exploiting in the customization exercise.

7. Sensitivity to Integer Constraints

The likelihood function for our model in Equation (12) assumes that consumers solve the utility maximization problem subject to nonnegativity and

Table 12 Sensitivity of Parameter Estimates and Elasticitites to Integer Constraints

Panel A: Parameter Estimates (Posterior Standard Deviations)

	$\varepsilon = 0$		$\varepsilon = 0.125$		$\varepsilon = 0.25$	
Flavor	$\bar{\psi}^*$	α^*	$\bar{\psi}^*$	α^*	$\bar{\psi}^*$	α^*
Strawberry	0.57 (0.01)	0.32 (0.06)	0.63 (0.01)	0.23 (0.05)	0.43 (0.03)	0.17 (0.09)
Blueberry	−0.15 (0.07)	0.66 (0.12)	−0.16 (0.13)	0.69 (0.26)	−0.39 (0.14)	−0.22 (0.21)
Piña Colada	0.12 (0.09)	0.02 (0.13)	0.18 (0.07)	−0.01 (0.11)	0.18 (0.06)	−0.57 (0.21)
Plain	−1.54 (0.12)	3.62 (0.50)	−1.49 (0.12)	3.77 (0.51)	−1.61 (0.12)	3.30 (0.53)
Mixed Berry	0.00	0.74 (0.11)	0.00	0.63 (0.12)	0.00	0.15 (0.14)

Panel B: Elasticities

Flavor	$\varepsilon = 0$	$\varepsilon = 0.125$	$\varepsilon = 0.25$
Strawberry	−2.21	−2.14	−1.82
Blueberry	−2.26	−2.30	−2.22
Piña Colada	−2.56	−2.50	−2.42
Plain	−1.88	−1.80	−1.26
Mixed Berry	−3.53	−3.57	−3.21

budget constraints. Our approach accommodates simultaneous purchases of more than one variety as well as corner solutions that are prominent in the data. However, we do not constrain the solutions to the utility maximization problem to consist of only integer quantities. In our data, consumers may only purchase yogurts in integer multiples of 8 ounces. It is well known that the evaluation of the likelihood for the fully integer-constrained problem is computationally extremely challenging as the mapping from the distribution of random utility errors to the quantity demanded is much more complicated than that provided by the standard Kuhn-Tucker conditions. To evaluate the likelihood for the integer-constrained problem, the probability of an observed n-tuple of integer demands much be computed by a complicated integral over a region in the random utility error space. Thus, our approach captures an important feature of the data (i.e., multiple discreteness) without the full integer solution. We contend that our approach provides a useful approximation to the full integer-constrained problem. In this section, we will conduct some sensitivity analysis that will shed light on the accuracy of this approximation.

The dominant feature of our data is not the purchase of multiple units of the same variety but zero demand for most varieties on most purchase occasions. It might be argued that the fact that our approach allows the consumer to select small positive quantities could distort the parameter estimates relative to the integer-constrained solution. That is, our noninteger-constrained approach will attempt to "fit" the discrete features of the data by selecting parameter estimates more extreme than the integer-constrained approach. In particular, the satiation or curvature parameters might be most affected. To assess the magnitude of these biases,[1] we reformulate the problem to regard a 0 in the data as arising from an optimal solution that is in the interval $(0, \varepsilon)$. This involves changing the limits of integration in the evaluation of the likelihood (see Equation (12)). We will investigate values of ε corresponding to 1 ounce and 2 ounces (0.125 and 0.25 because 1 unit is an 8 ounce container).

Table 12 provides the results of re-estimation of the inside goods model with various values of ε. For example, the middle column presents results for a model in which a 0 purchase quantity is interpreted

[1]We are grateful to a referee for making this suggestion.

Table 13 **Sensitivity of Compensating Value Computations to Integer Constraints**

Flavor Deleted	$\varepsilon = 0$	$\varepsilon = 0.125$	$\varepsilon = 0.25$
Strawberry	10.61	10.99	12.38
Blueberry	9.97	10.04	10.07
Piña Colada	9.19	9.71	10.19
Plain	14.27	15.13	18.72
Mixed Berry	8.48	8.69	8.98

as an optimal quantity demanded of 1 ounce or less. Panel A presents the common parameter estimates for baseline utility parameter and the satiation parameter, while Panel B provides the own price elasticity estimates for aggregated demand. Small perturbations of 1 ounce or less make no difference in either the parameter estimates or the elasticities. However, some differences arise at 2 ounces. In particular, the satiation parameters change to smaller values for all varieties, implying a higher level of curvature or increased satiation in our parameterization. This makes intuitive sense because the effect of nonzero values of ε is to increase the probabilities of corners. To "fit" the observed frequency of corners in our data, we must increase the curvature parameter. With increased satiation, the effect of a price change is less dramatic, so this accounts for the decline in elasticities.

Table 13 provides *CV* calculations using the set of parameters displayed in Table 12, Panel A. Utility and associated monetary equivalents is the right metric for assessing whether the parameters changes observed in Table 12, Panel A are substantively important. Table 13 shows that for $\varepsilon = 0.125$, we see no appreciable change in the *CV* calculations. However, for $\varepsilon = 0.25$, the *CV* for plain yogurt increases by 31% that is appreciable.

Thus, we conclude that imposition of full integer constraints has the possibility to change some of our results by economically meaningful amounts. It is important to emphasize that the calculations in Table 13 indicate that failure to impose full integer constraints reduces the utility loss from reduction in variety. This means that our overall conclusions about the high value for variety remain unchanged, as well as our conclusion that these utility losses are incurred primarily from a small fraction of households who highly value the deleted flavor.

8. Concluding Remarks

The observation of multiple purchases in the same product category rules out the use of standard linear utility-based choice models. Extensions of the standard multinomial choice models such as the multivariate probit model (cf. Manchanda et al. 1999) can accommodate multiple purchase data. However, these models are not based on a formal utility maximization framework and, therefore, cannot be used for pricing and product assortment policy decisions such as those considered in §5.

We propose a translated nonlinear additive utility function that accommodates mixtures of interior and corner solutions. The proposed function nests the commonly used linear utility structure as a special case, and can be used to measure rates of satiation of different offerings that give rise to the simultaneous purchase of multiple varieties. Further, the function provides a valid measure of consumer utility needed in considering the value of assortment variety. The likelihood function for this model is derived from an assumed distribution of marginal utility errors. Evaluation of this likelihood function requires the calculation of high-dimensional normal integrals, posing a barrier to the use of the class of translated utility models. Furthermore, we employ a Bayesian hierarchical approach to modeling household heterogeneity that allows us to make household-level utility calculations.

We use our model to undertake utility-based calculations that get at the value of variety. Our utility-based model provides an explicit metric for valuing variety derived from economic principles. Once calibrated, our utility metric can be used to assess counterfactual experiments such as evaluating the value of a reduced assortment. Large utility losses are incurred when a flavor is deleted from the assortment available to all households. These utility losses can be compensated for by substantial reduction in either the prices of remaining flavors or in the overall level of prices in the store. These large utility losses are driven by heterogeneity in preferences across households. Deletion of any one flavor hurts a subset of households. Uniform policy changes by the retailer (such as price reductions) are a rather crude and expensive way of compensating for the loss in variety. We compute the reduction in prices required to compensate for the loss in variety from deletion of a flavor. Large price reductions are required to keep the aggregate total utility-level constant between the full and limited assortments.

In nontraditional retailing contexts such as Web retailing, customization of the assortment presented to any one customer holds out the possibility of reducing variety without incurring substantial utility loss. Our customized assortment calculations show almost zero utility loss from deletion of one flavor as long as the flavor deleted is household-specific. The customization of assortment works well in the yogurt context because there are subsets of households with strong flavor preferences. As long as these households receive assortments that include their most preferred brand, they will suffer little utility loss from deletion of other flavors.

Horizontal variety is frequently observed to occur in product categories with high rates of consumption, packaged goods, and entertainment products. For less frequently consumed products, or for products in which small quantities are consumed at any point in time, evidence of variety seeking is accumulated temporally across purchase and/or consumption occasions. In our model, we include curvature parameters to measure satiation but we do not have a formal model of how satiation occurs. In particular, if marketing activities affect the timing or nature of the consumption occasion, then our satiation parameters would not be policy invariant. More formal dynamic modeling of satiation and variety seeking, as well as the way in which the marketing environment affects these behaviors, is a subject for future research.

In addition to the dynamic aspects of satiation, it would be useful to explore a characteristics model in which brand characteristics could explain the different satiation rates. For example, it is possible that the high satiation of the strawberry flavor may be due to factors such as the sweetness of the flavor. To investigate this issue, it is necessary to define utility on the characteristics of the goods and to regard the goods as bundles of these characteristics.

Acknowledgments
The authors thank participants at seminars at the University of Chicago, Northwestern University, New York University, Cornell University, University of Colorado, University of Toronto, and Stanford University for useful comments. In particular, we thank Mike Keane and Surendra Rajiv for helpful suggestions. Support from the Kilts Center for Marketing, Graduate School of Business, University of Chicago is gratefully acknowledged.

Appendix
MCMC Algorithm

A Metropolis-Hastings (M-H) algorithm is used within a "Gibbs" style MCMC sampler. Households ($h = 1, \ldots, H$) provided choice and quantity information (y_{ht}) for m alternatives at each occasion $t = 1, \ldots, T_h$. L_{ht} denotes the likelihood of a household (h)'s purchase for m alternatives at each occasion (t) expressed by Equation (7). The estimation algorithm proceeds by recursively generating draws from the following densities.

1. Generate $\{\boldsymbol{\psi}_h^*, h = 1, \ldots, H\}$

$$\pi(\boldsymbol{\psi}_h^* | \{y_{ht}, t = 1, \ldots, T_h\}, \boldsymbol{\alpha}, \bar{\boldsymbol{\psi}}*, \mathbf{D}_{\psi^*}) \propto \det|\mathbf{D}_{\psi^*}|^{-1/2} \exp\left[-\tfrac{1}{2}(\boldsymbol{\psi}_h^* - \bar{\boldsymbol{\psi}}^*)'\mathbf{D}_{\psi*}^{-1}(\boldsymbol{\psi}_h^* - \bar{\boldsymbol{\psi}}^*)\right] \cdot \prod_t^{T_h} L_{ht}.$$

To obtain a draw for the ($m - 1 \times 1$) vector $\boldsymbol{\psi}_h^*$, the M-H algorithm proceeds as follows. Let $\boldsymbol{\psi}_h^{*(k)}$ be the kth draw for $\boldsymbol{\psi}_h^*$. The next draw ($k + 1$) is given by

$$\boldsymbol{\psi}_h^{*(k+1)} = \boldsymbol{\psi}_h^{*(k)} + \boldsymbol{\delta}_\psi,$$

where $\boldsymbol{\delta}_\psi$ is a draw from candidate generating density N (0, 0.0025). The probability of accepting the new draw, $\boldsymbol{\psi}_h^{*(k+1)}$ is given by

$$\min\left[\frac{\exp\left[-\frac{1}{2}(\boldsymbol{\psi}_h^{*(k+1)} - \bar{\boldsymbol{\psi}}^*)'\mathbf{D}_{\psi^*}^{-1}(\boldsymbol{\psi}_h^{*(k+1)} - \bar{\boldsymbol{\psi}}^*)\right] \cdot \prod_t L_{ht}^{(k+1)}}{\exp\left[-\frac{1}{2}(\boldsymbol{\psi}_h^{*(k)} - \bar{\boldsymbol{\psi}}^*)'\mathbf{D}_{\psi*}^{-1}(\boldsymbol{\psi}_h^{*(k)} - \bar{\boldsymbol{\psi}}^*)\right] \cdot \prod_t L_{ht}^{(k)}}, 1\right].$$

2. Generate $\mathbf{D}_{\bar{\psi}^*}$

$$\pi(\bar{\boldsymbol{\psi}}^*|\{\boldsymbol{\psi}_h^*, h=1,\ldots,H\}, \mathbf{D}_{\psi^*}) = N\left(\frac{\sum_{h=1}^{H}\boldsymbol{\psi}_h^*}{H}, \frac{\mathbf{D}_{\psi*}}{H}\right).$$

3. Generate $\mathbf{D}_{\psi^*}$

$$\pi(\mathbf{D}_{\psi*}|\{\boldsymbol{\psi}_h^*, h=1,\ldots,H\}, \bar{\boldsymbol{\psi}}^*)$$
$$\propto \text{Inverted Wishart}\left(d_0 + H, D_0 + \sum_{h=1}^{H}(\boldsymbol{\psi}_h^* - \bar{\boldsymbol{\psi}}*)(\boldsymbol{\psi}_h^* - \bar{\boldsymbol{\psi}}^*)'\right),$$

where $\boldsymbol{d}_0$ and $\boldsymbol{D}_0$ the prior degrees of freedom and sum of squares for $\mathbf{D}_{\psi^*}$.

4. Generate $\boldsymbol{\alpha}^*$

$$\pi(\boldsymbol{\alpha}^*|\{\boldsymbol{y}_{ht}, h=1,\ldots,H \text{ and } t=1,\ldots,T_h\},$$
$$\{\boldsymbol{\psi}_h^*, h=1,\ldots,H\}, \bar{\boldsymbol{\alpha}}_0, \boldsymbol{\Sigma}_0)$$
$$\propto \det|\boldsymbol{\Sigma}_0|^{-1/2}\exp\left[-\tfrac{1}{2}(\boldsymbol{\alpha}^* - \bar{\boldsymbol{\alpha}}_0)'\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\alpha}* - \bar{\boldsymbol{\alpha}}_0)\right]\cdot\prod_{h=1}^{H}\prod_{t=1}^{T_h}L_{ht},$$

where $\bar{\boldsymbol{\alpha}}_0$ and $\boldsymbol{\Sigma}_0$ are prior parameters. In our M-H algorithm, the $(k+1)^{th}$ $\boldsymbol{\alpha}^*$ is given by

$$\boldsymbol{\alpha}^{*(k+1)} = \boldsymbol{\alpha}^{*(k)} + \boldsymbol{\delta}_a,$$

where $\boldsymbol{\delta}_a$ is a draw from candidate generating density N (0, 0.01). The probability of accepting the new draw, $\boldsymbol{\alpha}^{*(k+1)}$is given by

$$\min\left[\frac{\exp\left[-\frac{1}{2}(\boldsymbol{\alpha}^{*(k+1)} - \bar{\boldsymbol{\alpha}}_0)'\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\alpha}^{*(k+1)} - \bar{\boldsymbol{\alpha}}_0)\right]\cdot\prod_h\prod_t L_{ht}^{(k+1)}}{\exp\left[-\frac{1}{2}(\boldsymbol{\alpha}^{*(k)} - \bar{\boldsymbol{\alpha}}_0)'\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\alpha}^{*(k)} - \bar{\boldsymbol{\alpha}}_0)\right]\cdot\prod_h\prod_t L_{ht}^{(k)}}, 1\right].$$

References

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

Allenby, Greg M., ——. 1999. Marketing models of heterogeneity. *J. Econom.* **89** 57–78.

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

Berry, Steven, James Levinsohn, Ariel Pakes. 1995. Automobile prices in market equilibrium. *Econometrica* **63** (4) 841–890.

Besanko, D., S. Gupta, D. Jain. 1998. Logit demand estimation under competitive pricing behavior. *Management Sci.* **44** 1533–1547.

Bradlow, E., V. Rao. 2000. A hierarchical Bayes model for assortment choice. *J. Marketing Res.* **37** 259–268.

Chiang, Jeongwon. 1991. A simultaneous approach to whether to buy, what to buy, and how much to buy. *Marketing Sci.* **4** (Fall) 297–314.

Chib, Sid, Edward Greenberg. 1995. Understanding the Metropolis-Hastings algorithm. *Amer. Statistician* **49** (4) 327–335.

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

Deaton, A., J. Muelbauer. 1980. *Economics and Consumer Behavior.* Cambridge University Press, Cambridge, MA.

Dube, J. P. 1999. Multiple discreteness and the demand for carbonated soft drinks. Working paper, Northwestern University, Evanston, IL.

Erdem, T. 1996. Dynamic analysis of market structure based on panel data. *Marketing Sci.* **15** (4) (Fall) 203–238.

Hajivassiliou, V., D. McFadden, P. Ruud. 1996. Simulation of multivariate normal rectangle probabilities and their derivatives. *J. Econometrics* **72** 85–134.

Hanneman, Michael. 1984. The discrete/continuous model of consumer demand. *Econometrica* **52** 541–561.

Hendel, I. 1999. Estimating multiple-discrete choice models: An application to computerization returns. *Rev. of Econom. Stud.* **66** 423–446.

Kahn, B. E., M. Kalwani, D. Morrison. 1986. Measuring variety-seeking and reinforcement behaviors using panel data. *J. Marketing Res.* **23** (May) 89–100.

Keane, M. 1994. A computationally practical simulation estimator for panel data. *Econometrica* **62** 95–116.

Lattin, James M., Leigh McAlister. 1985. Using a variety-seeking model to identify substitute and complementary relationships among competing products. *J. Marketing Res.* **22** (August) 330–339.

Manchanda, P., A. Ansari, S. Gupta. 1999. The "shopping basket": A model for multicategory purchase incidence decisions. *Marketing Sci.* **18** 95–114.

McAlister, Leigh. 1982. A dynamic attribute satiation model of variety-seeking behavior. *J. Consumer Res.* **9** (September) 141–151.

——, Edgar Pessemier. 1982. Variety seeking behavior: An interdisciplinary review. *J. Consumer Res.* **9** (December) 311–322.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller. 1953. Equations of state calculations by fast computing machines. *J. Chemical Phys.* **21** 1087–1092.

Nevo, Aviv. 2001. Measuring market power in the ready-to-eat cereal industry. *Econometrica* **69** 307–342.

Rossi, Peter, Robert McCulloch, Greg Allenby. 1996. On the value of household information in target marketing. *Marketing Sci.* **15** 321–340.

Wales, T. J., A. D. Woodland. 1983. Estimation of consumer demand systems with binding non-negative constraints. *J. Econometrics* **21** 437–468.

Walsh, J. W. 1995. Flexibility in consumer purchasing for uncertain future tastes. *Marketing Sci.* **14** 148–165.

This paper was received December 29, 2000, and was with the authors 4 months for 3 revisions; processed by Dick Wittink.