Investigating Heterogeneity in Brand Preferences in Logit Models for Panel Data
Author(s): Pradeep K. Chintagunta, Dipak C. Jain and Naufel J. Vilcassim
Source: *Journal of Marketing Research*, Vol. 28, No. 4 (Nov., 1991), pp. 417-428
Published by: American Marketing Association
Stable URL: http://www.jstor.org/stable/3172782

PRADEEP K. CHINTAGUNTA, DIPAK C. JAIN, and NAUFEL J. VILCASSIM*

In analyzing panel data, the issue of heterogeneity across households is an important consideration. If heterogeneity is present but is ignored in the analysis, it will result in biased and inconsistent estimates of the effects of marketing mix variables on brand choice. The authors propose the use of a random effects specification to account for heterogeneity in brand preferences across households in a logit framework. The model parameters are estimated by both parametric and semiparametric approaches. The authors also compare their results with those obtained from logit models in which observed past choice behavioir is used to capture such heterogeneity. The different models are estimated with the IRI saltine crackers dataset. A formal statistical test of the model specifications shows that the semiparametric specification is the most preferred in terms of the overall fit of the model to the data. In addition, that specification predicts best when the models are validated in a holdout sample of households.

Investigating Heterogeneity in Brand Preferences in Logit Models for Panel Data

The conditional logit model developed by McFadden (1974) has been widely applied with panel data to analyze brand choice decisions of households (e.g., Gensch and Recker 1979; Guadagni and Little 1983; Gupta 1988; Krishnamurthi and Raj 1988; Zufryden 1986). The theoretical development of the logit model is based on utility-maximizing behavior at the individual or household level. Therefore, ideally, the parameters of the logit model should be estimated at the household level. In panel data, however, the number of observations per household is often insufficient for consistent and efficient estimation of household-specific parameters. Further, from the standpoint of marketing decision making, the parameter estimates are meaningful only at the aggregate or market level. Accordingly, researchers have resorted to pooling the data across households and estimating a set of aggregate-level parameters.

A pertinent question is: What is the effect of heterogeneity across households on the estimated parameters of a logit model of brand choice? In general, pooling across households while ignoring heterogeneity when it is present will lead to biased and inconsistent estimates of the effects of marketing variables (Hsiao 1986).

We investigate heterogeneity across households by proposing a parsimonious model that captures the variations across households in their overall brand preferences.[1] We also compare the proposed method with other methods used in the literature to account for heterogeneity in preferences. We determine the appropriateness of the different specifications for heterogeneity in preferences by using a formal statistical test and also by comparing the predictive validity of the different models in a holdout sample of households. Such a comparative

*Pradeep K. Chintagunta is Assistant Professor, S. C. Johnson Graduate School of Management, Cornell University. Dipak C. Jain is Associate Professor of Marketing and Naufel Vilcassim is Assistant Professor of Marketing, J. L. Kellogg Graduate School of Management, Northwestern University.

The authors thank Bo Honore, Rishin Roy, and Dick Wittink, three anonymous *JMR* reviewers, and the editor for their valuable suggestions and insightful comments. The financial support provided to the second author by the Kraft Research Professorship and to the third author by the Center for Marketing Sciences at the J. L. Kellogg Graduate School of Management, Northwestern University, is greatly appreciated.

[1]This heterogeneity across households has also been referred to as "unobserved heterogeneity" in the literature as it captures the effects of unobserved factors that influence household choice behavior (Jain and Vilcassim 1991; Lancaster 1979).

Journal of Marketing Research
Vol. XXVIII (November 1991), 417–28

analysis should be useful to researchers modeling brand choice behavior with logit models.

Previously researchers in marketing have attempted to incorporate heterogeneity in brand preferences in the context of a logit model using observed past brand choice behavior. For example, Guadagni and Little (1983) used two variables referred to as "brand loyalty" and "size loyalty," which are exponentially weighted averages of past brand and brand-size choices, to account for heterogeneity across households. Krishnamurthi and Raj (1988) used a household-specific variable based on the share of purchases of a particular brand in relation to all brands. In both cases, the value of the heterogeneity variable changes from one purchase occasion to the next. More recently, Allenby and Rossi (1990) used a somewhat different operationalization, wherein heterogeneity is captured via a fixed term that is measured as the difference between the predicted choice probabilities and the probabilities estimated by a relative frequency approach. A justification for the use of such variables is that they attempt to capture the differences in brand preferences across households.

A question that arises in using observed past brand choice behavior to capture heterogeneity across households is whether it would affect the estimates of the marketing mix variables such as price and promotional activities. Srinivasan and Kibarian (1990) claim that the effect of past brand choices on current choice will be overstated if the past choices are not adjusted for the possible effects of price and other promotional variables. Consequently, the effect of some of the other variables included in the model would be understated.

An alternative approach to capturing heterogeneity in preferences across households is to include a household-specific fixed or constant term in addition to a common set of parameters for the marketing mix variables. The rationale is that the preference of a household for each brand is stable and can be captured adequately by a fixed term. Such an approach to capturing the unobserved difference in preferences across households is not dependent on past choice behavior and therefore should not contaminate the estimated effects of the marketing mix variables.

There are two approaches to estimating the model parameters with such a specification. One is to estimate the constant term for each household for each brand, which is referred to as a "fixed-effects" model in the literature (Chamberlain 1980). However, it involves estimating a large number of parameters and requires long purchase strings for each household.[2] When there are only a few observations for each household, conventional maximum likelihood estimation will lead to inconsistent estimates not only of the fixed term, but also of the effects of marketing mix variables (Hsiao 1986).

To circumvent that problem, Chamberlain (1980) suggests the method of conditional maximum likelihood estimation (CMLE) to obtain consistent estimates for the effects of the included marketing variables. CMLE involves conditioning the likelihood function on sufficient statistics for the fixed household-specific terms. This approach renders the conditional likelihood function independent of the fixed terms, thereby providing consistent estimates for the other parameters in the model. Such an estimation procedure has been used by Jones and Landwehr (1988) and also by Kudpi (1989) for the case of two choice alternatives. When the number of alternatives is large, the CMLE procedure becomes extremely complex as the conditioning involves a large number of possible permutations of the purchase sequences of various brands. Further, in the CMLE one does not estimate the fixed term for each household, and hence such a model cannot be used to predict household brand choice probabilities.

A more tractable approach to estimating the model parameters is to assume that the fixed term varies across households according to some probability distribution. This method of accounting for heterogeneity is referred to as a "random effects" specification (Heckman 1981; Hsiao 1986). The model parameters under such a specification can be estimated in two ways. One is to assume a particular parametric form for the distribution of heterogeneity across households. For example, one could assume a gamma distribution for the heterogeneity and estimate for each brand the associated parameters, in addition to a common set of parameters for the marketing mix variables. Such a parametric approach would be consistent with the manner in which heterogeneity has been accounted for in applied stochastic models of household purchase behavior (e.g., Bass 1974; Bass, Jeuland, and Wright 1976). A potential problem is that if one imposes an incorrect probability distribution for the heterogeneity across panelists, it would result in biased estimates of the effects of marketing mix variables on brand choice (Heckman and Singer 1984).

A solution to that problem is to use a semiparametric random effects specification in which one does not impose any parametric form for the heterogeneity across households, but instead estimates the distribution empirically using the data. The strategy is to approximate the underlying probability distribution across households for each brand by a finite number of support points. The estimation involves determining the location of those points and the probability masses associated with each support point. This method has been used recently by Heckman and Singer (1984), Gonul and Srinivasan (1990), and Vilcassim and Jain (1991).

We propose the use of a random effects specification for the heterogeneity across households and estimate the model parameters using both parametric and semiparametric approaches. In the parametric case, we assume

[2]For example, if the panel consists of 100 households and four brands in the product category, it would involve estimating 300 (= 100 * (4 − 1)) fixed-effect terms.

the heterogeneity to follow either a gamma or a normal distribution across households, as those distributions have been widely used in the marketing literature (e.g., Bass, Jeuland, and Wright 1976; Jain and Vilcassim 1991). A potential difficulty with the use of those distributions is that they do not lead to closed-form expressions for the brand choice probabilities. Though some numerical integration technique can be used to estimate the model parameters, it is often extremely difficult in practice and leads to unstable results. We show how such estimation problems can be overcome by using a Gaussian quadrature (Press et al. 1986) to approximate the integrals in the expressions for the choice probabilities.

We also estimate the distribution of heterogeneity semiparametrically and compare the results with those obtained from using a parametric specification. For completeness, we also compare our results with those obtained from three additional logit models of brand choice using the Guadagni and Little (1983), Krishnamurthi and Raj (1988), and Allenby and Rossi (1990) operationalizations of heterogeneity described previously. As it would be useful to know which of these six different specifications of heterogeneity in preferences is most consistent with the data,[3] we use two different criteria to choose among them, (1) a formal statistical test and (2) predictive validity of the models in a holdout sample of households. On the basis of those two criteria, we identify the preferred specification of heterogeneity and discuss its substantive marketing implications.

In a recent study, Colombo and Landwehr (1990) also examined how heterogeneity in brand preferences can be estimated by a semiparametric approach. The distinguishing feature of our study is that we provide a comprehensive empirical analysis by comparing the proposed semiparametric approach with (1) alternative parametric random effects specifications and (2) approaches that have used past brand choice behavior to account for heterogeneity. Such a comparative analysis should be helpful to marketing researchers using logit models to analyze household brand choice behavior.

Our empirical results from analyzing the IRI saltine crackers data suggest that the semiparametric random effects specification provides the best overall fit to the data. This finding is confirmed by a formal statistical test of logit models that are mutually non-nested. Further, the specification also predicts best in the holdout sample of households, thereby suggesting that the semiparametric random effects specification is the preferred method of accounting for heterogeneity in preferences across households in logit models of brand choice.

In the next section, we describe the specification of a random effects logit model and discuss issues relating to the estimation of the model parameters. We then discuss the marketing implications of the empirical results obtained from analyzing the IRI household panel data on purchases of saltine crackers. We summarize the findings in the last section, discuss the limitations of the study, and provide directions for future research.

[3]The six specifications are: two parametric (gamma and normal) specifications, one semiparametric specification, and three operationalizations of the brand loyalty variable.

MODEL SPECIFICATION AND ESTIMATION

Consider a sample of M panelists $(1, 2, \ldots, i, \ldots, M)$ and a choice set of N brands $(1, 2, \ldots, j, \ldots, N)$. Following McFadden (1974), we assume that the indirect utility U_{ijt} of the i^{th} panelist for brand j on choice occasion t consists of a deterministic component V_{ijt} and a random component ϵ_{ijt}:

$$U_{ijt} = V_{ijt} + \epsilon_{ijt}.$$

The deterministic component V_{ijt} consists of a systematic part $\mathbf{X}_{ijt}\boldsymbol{\beta}$ (where $\mathbf{X}_{ijt}$ is a vector of marketing mix variables or covariates confronting the panelist i on choice occasion t and $\boldsymbol{\beta}$ is a vector of unknown parameters) and a fixed term α_{ij}, which captures the intrinsic preference of panelist i for brand j:

$$V_{ijt} = \mathbf{X}_{ijt}\boldsymbol{\beta} + \alpha_{ij}.$$

Under the assumptions of utility maximization and an independent and identically distributed type 1 extreme value distribution for ϵ_{ijt}, the probability that panelist i chooses brand j on choice occasion t (see, e.g., Ben-Akiva and Lerman 1985) conditional on $\alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{ij}, \ldots, \alpha_{iN}$ is given by

$$(1) \qquad Pr_{it}(j \mid \alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{iN}) = \frac{\exp(\mathbf{X}_{ijt}\boldsymbol{\beta} + \alpha_{ij})}{\sum_{k=1}^{N} \exp(\mathbf{X}_{ikt}\boldsymbol{\beta} + \alpha_{ik})}.$$

If the covariate effects are the same across all brands, the brand choice decision is determined by the α_{ik} values. Hence, we can treat the α_{ik} term as representing the intrinsic preference of panelist i for brand k. For a given household i, α_{ik}'s are assumed to be N fixed scalars that do not vary over time. Across households, however, they may vary. We attempt to capture these variations in the α_{ik}'s (i.e., heterogeneity) by assuming a probability distribution for each brand. Let α_k denote the random variable associated with the probability distribution for brand k, $k = 1, 2, \ldots, N$. Therefore, α_{ik} represents the particular realization of α_k for household i. We treat the distribution of preferences across households for a particular brand as independent of the distribution of preferences for the other brands.

To identify the model parameters, we normalize equation 1 with respect to some base brand, N (say).[4] Then,

[4]We note that when $\mathbf{X}_{ijt}$ includes household-specific, time-invariant variables (e.g., income, family size), such variables will drop out after normalization in equation 2. In that case, it is necessary to make the coefficients of such variables brand specific (see Allenby and Rossi 1990; Krishnamurthi and Raj 1988).

$$(2)\quad Pr_{it}(j \mid \alpha_{i1}, \alpha_{i2}, \ldots, \alpha_{iN}) = \frac{\exp\{(\mathbf{X}_{ijt} - \mathbf{X}_{iNt})\boldsymbol{\beta} + (\alpha_{ij} - \alpha_{iN})\}}{1 + \sum_{k=1}^{N-1} \exp\{(\mathbf{X}_{ikt} - \mathbf{X}_{iNt})\boldsymbol{\beta} + (\alpha_{ik} - \alpha_{iN})\}} = \frac{\exp(\mathbf{Z}_{ijt}\boldsymbol{\beta} + \nu_{ij})}{1 + \sum_{k=1}^{N-1} \exp(\mathbf{Z}_{ikt}\boldsymbol{\beta} + \nu_{ik})}$$

where $\mathbf{Z}_{ijt} = (\mathbf{X}_{ijt} - \mathbf{X}_{iNt})$ and $\nu_{ik} = \alpha_{ik} - \alpha_{iN}$, $k = 1, 2, \ldots, j, \ldots, N - 1$.

Equation 2 represents the conditional probability that household i will purchase brand j given the intrinsic preferences, α_{ik}, $k = 1, 2, \ldots, N$. In general, because of the insufficient number of observations on the purchase behavior of household i, it would not be feasible to estimate the household-specific parameter ν_{ik}, $k = 1, 2, \ldots, N - 1$. To circumvent that problem, we derive the unconditional probability that a randomly selected household will choose brand j. That probability could be obtained by integrating the conditional probability in equation 2 with respect to the probability distribution of ν_k, $k = 1, 2, \ldots, N - 1$, where $\nu_k = \alpha_k - \alpha_N$. We note that though the random variables α_k, $k = 1, 2, \ldots, N$, are assumed to be independent, the ν_k's are not independently distributed because

$$E[\nu_k, \nu_l] = E[(\alpha_k - \alpha_N)(\alpha_l - \alpha_N)], \qquad l \neq k$$
$$= E[(\alpha_N)^2] \neq 0.$$

Hence, working with the random variables ν_k, $k = 1, 2, \ldots, N - 1$, requires the use of an $(N - 1)$-variate joint probability distribution, which makes the estimation procedure extremely difficult, if not intractable.

To keep the estimation tractable, we propose the following decomposition approach that exploits the assumption of independence of the random variables α_k, $k = 1, 2, \ldots, N$. We decompose the random variable α_k as follows:

$$(3)\qquad \alpha_k = \mu_k + \gamma_k, \qquad k = 1, 2, \ldots, N,$$

where μ_k is a random variable with mean zero and γ_k is a brand-specific scalar quantity invariant across households and equal to the mean value of α_k. As the random variables α_k are independent, equation 3 implies that the random variables μ_k, $k = 1, 2, \ldots, N$, are also independent. Accordingly, for the i^{th} panelist we have

$$(4)\qquad \alpha_{ik} = \mu_{ik} + \gamma_k, \qquad k = 1, 2, \ldots, N.$$

Equation 1 for the conditional choice probability then can be written as

$$(5)\quad Pr_{it}(j \mid \mu_{i1}, \ldots, \mu_{iN}) = \frac{\exp(\mathbf{X}_{ijt}\boldsymbol{\beta} + \mu_{ij} + \gamma_j)}{\sum_{k=1}^{N} \exp(\mathbf{X}_{ikt}\boldsymbol{\beta} + \mu_{ik} + \gamma_k)}.$$

In normalizing equation 5, an important consideration is to be able to preserve the assumption of independence of the random variables μ_k, $k = 1, 2, \ldots, N$. Accordingly, we divide by $\exp(\mathbf{X}_{iNt}\boldsymbol{\beta} + \gamma_N)$ and obtain

$$(6)\quad Pr_{it}(j \mid \mu_{i1}, \ldots, \mu_{iN}) = \frac{\exp\{(\mathbf{X}_{ijt} - \mathbf{X}_{iNt})\boldsymbol{\beta} + \mu_{ij} + (\gamma_j - \gamma_N)\}}{\exp(\mu_{iN}) + \sum_{k=1}^{N-1} \exp\{(\mathbf{X}_{ijt} - \mathbf{X}_{iNt})\boldsymbol{\beta} + \mu_{ik} + (\gamma_k - \gamma_N)\}}.$$

Setting $\mathbf{Z}_{ijt} = \mathbf{X}_{ijt} - \mathbf{X}_{iNt}$, and $\theta_j = \gamma_j - \gamma_N$, $j = 1, 2, \ldots, N$, we obtain

$$(7)\quad Pr_{it}(j \mid \mu_{i1}, \ldots, \mu_{iN}) = \frac{\exp(\mathbf{Z}_{ijt}\boldsymbol{\beta} + \mu_{ij} + \theta_j)}{\exp(\mu_{iN}) + \sum_{k=1}^{N-1} \exp(\mathbf{Z}_{ijk}\boldsymbol{\beta} + \theta_k)}.$$

Let $\xi_{ik} = \exp(\mu_{ik})$, $k = 1, 2, \ldots, N$. The conditional probability in equation 7 can then be written as

$$(8)\quad Pr_{it}(j \mid \xi_{i1}, \ldots, \xi_{iN}) = \frac{\xi_{ij} \exp(\mathbf{Z}_{ijt}\boldsymbol{\beta} + \theta_j)}{\xi_{iN} + \sum_{k=1}^{N-1} \xi_{ik} \exp(\mathbf{Z}_{ikt}\boldsymbol{\beta} + \theta_k)}.$$

The heterogeneity across panelists now can be captured by the probability distribution of the random variables ξ_k, $k = 1, 2, \ldots, N$. By excluding ξ_{iN} in the normalization, we preserve the independence property of ξ_k, $k = 1, 2, \ldots, N$. This leads to tractability in the estimation. Further, the parameters θ_k, $k = 1, 2, \ldots, N - 1$, represent the mean intrinsic preference for each brand in relation to brand N, and therefore are analogous to the brand-specific constants used in the conventional logit models (e.g., Guadagni and Little 1983).

To obtain the unconditional probability of choosing brand j on purchase occasion t, we integrate equation 8 with respect to *all* N independent random variables ξ_1, $\xi_2, \ldots, \xi_N$. For notational convenience, we drop the subscript i in equation 8 and obtain

$$(9)\quad Pr_t(j) = \int_0^\infty \int_0^\infty \cdots \int_0^\infty Pr_t(j \mid \xi_1, \xi_2, \ldots, \xi_N) \cdot f(\xi_1) f(\xi_2) \ldots f(\xi_N)\, d\xi_1 d\xi_2 \ldots d\xi_N,$$

where $f(\xi_k)$ is the density function of ξ_k, $k = 1, 2, \ldots, N$. We have assumed here that the distribution of intrinsic preferences is independent of the marketing mix variables. Substituting equation 8 for $Pr_t(j \mid \xi_1, \ldots, \xi_N)$ in equation 9 yields

$$(10)\quad Pr_t(j) = \int_0^\infty \cdots \int_0^\infty \int_0^\infty \cdot \left\{ \frac{\xi_j \exp(\mathbf{Z}_{jt}\boldsymbol{\beta} + \theta_j)}{\xi_N + \sum_{k=1}^{N-1} \xi_k \exp(\mathbf{Z}_{kt}\boldsymbol{\beta} + \theta_k)} \right\} \cdot f(\xi_1) f(\xi_2) \ldots f(\xi_N)\, d\xi_1 d\xi_2 \ldots d\xi_N.$$

We note that equation 10 represents an aggregate prediction of the choice probability of brand j, that is, the probability that a randomly drawn household will choose brand j on occasion t. Further, we find that the ratio of $Pr_t(j)$ to $Pr_t(l)$, $j \neq l$, depends on *all* N brands in the choice set. Accordingly, the aggregate logit model (equation 10) that uses a random effects specification of heterogeneity across households belongs to the class of flexible discrete choice models discussed by Dalal and Klein (1988). The major advantage of this class of models is that the aggregate predicted brand choice probabilities from such models would be close (in magnitude) to the probabilities obtained from a random utility model derived under any arbitrary distribution of utilities. Hence, equation 10 represents a robust specification of the aggregate brand choice probabilities.

Two measures of managerial interest in the context of a logit model of brand choice are the aggregate own- and cross-choice elasticities. For the random effects specification, the aggregate own-choice elasticity E^l_{jj} with respect to the l^{th} covariate of the brand j is given by

$$E^l_{jj} = \frac{\sum_{i=1}^{M}\sum_{t=1}^{T_i}\left\{\frac{\partial Pr_t(j)}{\partial \mathbf{X}_{ljt}} \cdot \frac{\mathbf{X}_{ljt}}{Pr_t(j)}\right\}}{\sum_{i=1}^{M} T_i}.$$

We note that in the preceding expression the value of the covariate $\mathbf{X}_{ljt}$ will vary across households. Substituting equation 9 for $Pr_t(j)$, we get after simplifying

$$E^l_{jj} = \frac{\sum_{i=1}^{M}\sum_{t=1}^{T_i}}{\sum_{i=1}^{M} T_i}\left\{\frac{\boldsymbol{\beta}_l \mathbf{X}_{ljt}}{Pr_t(j)}\right\} \cdot \left\{\left[\int_0^\infty \cdots \int_0^\infty Pr_t(j|\xi_1, \xi_2, \ldots, \xi_N) \cdot (1 - Pr_t(j|\xi_1, \xi_2, \ldots, \xi_N)) f(\xi_1)\ldots f(\xi_N) d\xi_1 \ldots d\xi_N\right]\right\}$$

where $\boldsymbol{\beta}_l$ is the coefficient associated with the l^{th} covariate. Similarly, the cross-choice elasticity E^l_{jk} of brand k with respect to the l^{th} covariate of brand j is given by

$$E^l_{jk} = \frac{\sum_{i=1}^{M}\sum_{t=1}^{T_i}\left\{\frac{\partial Pr_t(k)}{\partial \mathbf{X}_{ljt}} \cdot \frac{\mathbf{X}_{ljt}}{Pr_t(k)}\right\}}{\sum_{i=1}^{M} T_i} = \frac{\sum_{i=1}^{M}\sum_{t=i}^{T_i}}{\sum_{i=1}^{M} T_i} \cdot \left\{\frac{\boldsymbol{\beta}_l \mathbf{X}_{ljt}}{Pr_t(k)}\right\} \cdot \left\{\left[\int_0^\infty \cdots \int_0^\infty Pr_t(j|\xi_1, \xi_2, \ldots, \xi_N) Pr_t(k|\xi_1, \xi_2, \ldots, \xi_N) \cdot f(\xi_1), f(\xi_2), \ldots, f(\xi_N) d\xi_1, d\xi_2, \ldots, d\xi_N\right]\right\}.$$

The preceding expressions for the elasticities imply asymmetric or differential patterns of competition among brands, which are a consequence of aggregating probabilities with nonlinear functional forms. Such differential patterns of competition seem to characterize many frequently purchased product categories (Blattberg and Wisniewski 1989).

Estimation Issues

We use the method of maximum likelihood to estimate the model parameters, which consist of $\boldsymbol{\beta}$, the effects of the marketing mix variables; θ_k, $k = 1, 2, \ldots, N - 1$, the mean intrinsic brand preferences; and the parameters of the probability distribution of ξ_k, $k = 1, 2, \ldots, N$. Let T_i denote the number of brand purchases made by the i^{th} panelist. The likelihood function for the associated string of purchases, conditional on $\xi_1, \xi_2, \ldots, \xi_N$, is given by

$$(11) \qquad L_i|\xi_1, \ldots, \xi_N = \prod_{t=1}^{T_i}\left\{\prod_{j=1}^{N} Pr_t(j|\xi_1, \xi_2, \ldots, \xi_N)^{\delta_{ijt}}\right\}$$

where:

$$\delta_{ijt} = \begin{cases} 1 & \text{if the } i^{\text{th}} \text{ panelist chose brand } j \text{ on purchase occasion } t, \\ 0 & \text{otherwise.} \end{cases}$$

The unconditional likelihood function is obtained by integrating equation 11 over the probability distributions of $\xi_1, \xi_2, \ldots, \xi_N$:

$$(12) \qquad L_i = \int_0^\infty \cdots \int_0^\infty (L_i|\xi_1, \ldots, \xi_N) f(\xi_1) f(\xi_2), \ldots, f(\xi_N) d\xi_1, d\xi_2, \ldots, d\xi_N.$$

The sample likelihood function for the M panelists is

$$(13) \qquad L = \prod_{i=1}^{M} L_i = \prod_{i=1}^{M}\int_0^\infty \cdots \int_0^\infty \left\{\prod_{t=1}^{T_i}\left\{\prod_{j=1}^{N}\left[\frac{\xi_j \exp(\mathbf{Z}_{jit}\boldsymbol{\beta} + \theta_j)}{\xi_N + \sum_{k=1}^{N-1} \xi_k \exp(\mathbf{Z}_{kit}\boldsymbol{\beta} + \theta_k)}\right]^{\delta_{ijt}}\right\}\right\} f(\xi_1) f(\xi_2) \ldots f(\xi_N) d\xi_1 d\xi_2 \ldots d\xi_N.$$

To maximize the preceding sample likelihood function, we need to specify the probability distribution functions $f(\xi_1), f(\xi_2), \ldots, f(\xi_N)$. For the parametric random effects specification, we consider two candidate probability distributions, gamma and normal, which have been used previously in the marketing literature to model heterogeneity. In the case of the gamma distribution, the random variable is required to be defined over the range $(0, \infty)$. We note that $\xi_k = \exp(\mu_k)$ satisfies that condition. We therefore assume that the random variables ξ_k, $k = 1, 2, \ldots, N$, are gamma distributed with shape and scale parameters w_k and δ_k, respectively:

$$f(\xi_k) = \frac{\delta_k^{w_k}}{\Gamma(w_k)} \exp[-\delta_k \xi_k] \xi_k^{w_k - 1} \tag{14}$$

where $\Gamma(\cdot)$ is the gamma function.

Recall from equation 3 that μ_k is defined as a random variable with mean zero due to the presence of the brand-specific constant γ_k. As $\log(\xi_k) = \mu_k$, the random variable $\log(\xi_k)$ also has mean zero, which in turn imposes a functional relationship between the parameters w_k and δ_k of the distribution of ξ_k. Hence, we fix without loss of generality the parameter δ_k equal to 1, and estimate only the shape parameter (for details see Hausman, Hall, and Griliches 1984).

When the normal distribution is used to account for heterogeneity in mean preferences, the underlying random variable should be free to take both positive and negative values. The random variable ξ_k is not appropriate because it takes only non-negative values. Instead, we consider the random variable μ_k (see equation 3), which can take both positive and negative values. Note that μ_k has mean zero because of the presence of brand-specific constants γ_k. The density function for μ_k is

$$g(\mu_k) = \frac{1}{\sigma_k \sqrt{22}} \exp\left(-\frac{1}{2}\frac{\mu_k^2}{\sigma_k^2}\right). \tag{15}$$

Accordingly, the probability that a randomly drawn household will purchase brand j on occasion t is given by

$$Pr_t(j) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} Pr_t(j|\mu_1, \ldots, \mu_N) g(\mu_1) g(\mu_2) \ldots g(\mu_N) \, d\mu_1, d\mu_2, \ldots, d\mu_N.$$

When the expressions for the probability density functions of the gamma and standard normal distributions are substituted in equation 13, the resulting sample likelihood has no closed-form expression. Using numerical integration techniques is extremely cumbersome and often leads to unstable estimates. To avoid that problem, we use a Gaussian quadrature (Press et al. 1986) to approximate the integrals in the likelihood function by a summation of a finite number of points. The primary advantage of this method over a numerical integration technique is that it requires very few points (it turned out to be two in our case) to obtain a good approximation of the integrals in equation 13. We provide the details of the Gaussian quadrature method in the Appendix.

For the semiparametric random effects specification, we do not impose any parametric form for the probability density functions $f(\cdot)$ and $g(\cdot)$. Instead, we approximate each of the underlying probability distributions by a finite number of support points and determine the location and probability mass associated with each support point. The number of support points required to approximate each of the underlying probability distributions is determined empirically by an iterative procedure wherein support points are added until any two points become so close to each other that they cannot be distinguished empirically. The sample likelihood function for this specification is

$$L = \prod_{i=1}^{M} \left\{ \sum_{h=1}^{H_N} \sum_{h=1}^{H_{N-1}} \cdots \left\{ \sum_{h=1}^{H_1} \left\{ \prod_{t=1}^{T_i} \left\{ \frac{\xi_j^h \exp(\mathbf{Z}_{jit}\boldsymbol{\beta} + \theta_j)}{\xi_N^h + \sum_{k=1}^{N-1} \xi_k^h \exp(\mathbf{Z}_{kit}\boldsymbol{\beta} + \theta_k)} \right\}^{\delta_{ijt}} \right\} \lambda_1^h \right\} \ldots \lambda_N^h \right\}, \tag{16}$$

where ξ_j^h and λ_j^h are, respectively, the location and the mass associated with the h^{th} ($h = 1, 2, \ldots H_j$) support point for brand j. In expression 16, the random variables ξ_k, $k = 1, 2, \ldots, N$, are such that $E[\log \xi_k] = 0$, $k = 1, 2, \ldots, N$. We impose this restriction in estimating both the parametric and semiparametric specifications of heterogeneity in brand preferences.

DATA AND RESULTS

For the empirical estimation and predictive validation, we used the purchase data for a randomly selected sample of 135 panelists from the IRI saltine crackers database. We randomly split the data into estimation and holdout samples of 70 and 65 panelists, respectively. The estimation sample consisted of 1736 purchases and the holdout sample 1542 purchases. The product subcategory consists of four brands—Sunshine, Keebler, Nabisco, and private label. The last brand is a composite of private label offerings from different stores. The 16-ounce package was by far the most frequently purchased and only that package size was used. There were no multiple purchases on any given occasion for the panelists in the sample. The marketing mix variables used in the analysis were price, special display, and feature advertisement. The price variable was measured as the net purchase price, whereas dummy variables (coded 0,1) were used to represent the absence/presence of a special display and feature advertisement at the time of purchase.

The estimation was done by using the GAUSS programming language on a personal computer.[5] The pa-

[5]For the parametric specifications, we find that two points are adequate to approximate the underlying distribution of intrinsic preferences. This finding is consistent with those of Butler and Moffitt (1982). We also performed a sensitivity analysis using a 3-point approxima-

Table 1
PARAMETER ESTIMATES AND THEIR STANDARD ERRORS

Variable	Guadagni and Little[a]	Krishnamurthi and Raj	Allenby and Rossi[a]	Parametric (gamma)	Parametric (normal)	Semi-parametric
Display	.366	.376	.268	.384	.497	.546
	(.128)	(.132)	(.142)	(.177)	(.193)	(.183)
Feature	1.071	1.048	.750	.883	.974	.903
	(.190)	(.191)	(.213)	(.127)	(.140)	(.159)
Price	−3.964	−4.085	−4.138	−4.276	−5.131	−5.630
	(.414)	(.409)	(.461)	(.094)	(.088)	(.183)
Loyalty	3.570	3.532	3.345	—	—	—
	(.115)	(.115)	(.121)			
Constant, Sunshine[b]	.424	.475	.393	−.902	−.746	−.328
	(.165)	(.167)	(.202)	(.152)	(.177)	(.160)
Constant, Keebler	1.148	1.223	2.454	1.059	.816	.933
	(.217)	(.215)	(.269)	(.142)	(.142)	(.192)
Constant, Nabisco	1.640	1.772	1.781	1.012	2.232	4.527
	(.193)	(.192)	(.223)	(.178)	(.201)	(.211)
Log-likelihood	−878.42	−892.30	−897.51	−960.01	−876.68	−838.49
d.f.	1729	1729	1729	1726	1726	1710
$\bar{\rho}^{2}$[c]	.533	.524	.522	.488	.532	.546

[a]We used five observations to develop the loyalty variable.
[b]The normalization in the estimation was with respect to the private label brand.
[c]Calculated after adjustment for the number of parameters in each specification. The log-likelihood value for the model with only brand-specific constants was −1894.00.

rameter estimates and their standard errors for the six specifications are reported in Table 1. *Prima facie,* the results of the estimation appear to be similar for the six specifications. For example, all specifications produce results that have reasonable face validity (i.e., the estimated parameter values of the marketing mix variables are all significant at the 5% level and are of the expected signs). There are, however, some important differences.

First, in terms of the overall fit of the models, we see from Table 1 that the semiparametric random effects specification seems to be the best as measured by the log-likelihood or $\bar{\rho}^2$ values.[6] Second, the magnitudes of the parameter estimates vary across the different specifications, the difference in the estimated effect of the price variable being most striking.

These differences suggest the need for a formal statistical test to choose the specification that is most consistent with the data. A conventional likelihood ratio test of model specification is not appropriate as the models are mutually non-nested. Ben-Akiva and Lerman (1985) suggest the use of a non-nested test that is based on the computed $\bar{\rho}^2$ values (for details of the test, see Horowitz 1980). In that test, one model is treated as the null and the other models are tested, in turn, against the null model. The test statistic is given by

$$Pr(\bar{\rho}_2^2 - \bar{\rho}_1^2 > \tau) \leq \Phi\{-[-2\tau L(0) + (\eta_2 - \eta_1)]^{1/2}\}, \qquad \tau \geq 0,$$

where η_1 and η_2 are the number of parameters estimated for the two models and Φ denotes the cumulative distribution function of the normal distribution. The expression on the right side is an upper bound for the probability that the difference in the $\bar{\rho}^2$ values for the null and alternative models is greater than τ.

We note from Table 1 that among the three random effects specifications, the semiparametric approach has the largest $\bar{\rho}^2$ value (= .546). Likewise, the Guadagni and Little specification has the largest $\bar{\rho}^2$ value (= .533) among the other three specifications. Hence, we focus on determining which of those two specifications is most consistent with the data. The issue is whether the observed difference of .013 in the $\bar{\rho}^2$ values for the two models is statistically significant. Substituting the two values of $\bar{\rho}^2$, we obtain

$$Pr(\bar{\rho}_2^2 - \bar{\rho}_1^2 > .013) \leq \Phi\{-8.02\}.$$

The probability that this difference could have occurred by chance is less than .0001. Hence, we conclude that the semiparametric specification provides the best fit to the data.[7]

Another criterion for choosing between different model specifications is to evaluate them on their predictive capabilities in a holdout sample of households. We per-

tion and found that the improvement in the log-likelihood was very small (e.g., for the normal specification, the log-likelihood value increased from −876.68 to only −875.51).

[6] $\bar{\rho}^2$ is defined as

$$\bar{\rho}^2 = 1 - \frac{L - \Theta}{L(0)},$$

where L is the log-likelihood of the model being estimated, Θ is the number of parameters, and $L(0)$ is the log-likelihood of the model with only brand-specific constants (see Ben-Akiva and Lerman 1985). Hence, $\bar{\rho}^2$ adjusts for the number of parameters in each model.

[7]Note that one could also use other tests of non-nested hypotheses—for example, a Bayesian information criterion test as suggested by Allenby (1990).

form the predictive validation in the following steps. First, we compare the predictions in the entire holdout sample. Next, we do the comparison by dividing the holdout sample into two parts, one characterized by a high degree of repeat purchase and the other by a high degree of brand switching. The objective is to determine how robust the model specification is to the nature of household brand choice behavior. We noted previously that the semiparametric random effects specification is the best-fitting model of the three random effects approaches to accounting for heterogeneity. We therefore compare its predictive validity with that of the other three specifications that use past brand choice behavior to account for heterogeneity. The results of the predictive validation are given in Table 2.

We see from Table 2 that the semiparametric specification predicts better than the Guadagni and Little model because the percentages of brand choices correctly predicted are 78% and 71%, respectively. When the holdout sample is divided as discussed previously, we see from Table 2 that though the two specifications perform about equally in predicting brand choice under a high degree of repeat purchase (80% vs. 81%, respectively), the semiparametric specification predicts significantly better than the Guadagni and Little model in the sample characterized by a high degree of brand switching (74% vs. 48%). We also note for the latter sample that the Allenby and Rossi (1990) model performs better than the Guadagni and Little (1983) model. An explanation for this result is that the Guadagni and Little (1983) loyalty parameter is a function of the price and promotional activities of the brands purchased by the household. Promotion-sensitive households with moderate loyalty often switch when brands are promoted. Therefore, a loyalty measure based solely on past purchases will be confounded by promotional activity for such households. The Allenby and Rossi (1990) measure of loyalty, in contrast, uses model residuals to obtain a measure orthogonal to the price and promotion variables and would therefore perform better when there are more switchers in the sample. Given the predictions of the four models in Table 2, we conclude that overall the semiparametric random effects specification is the preferred one.

Table 2
MODEL VALIDATION RESULTS[a]

Brand	Actual share	Semi-parametric	GL	KR	AR
Loyals					
Sunshine	.02	.04	.03	.03	.03
Keebler	.04	.05	.03	.03	.03
Nabisco	.65	.64	.63	.63	.62
Private label	.29	.27	.31	.31	.32
RMSE[b]	—	.03	.03	.03	.04
Hit ratio (N = 1056)	—	80%	81%	81%	74%
Switchers					
Sunshine	.20	.18	.14	.15	.19
Keebler	.10	.12	.11	.11	.14
Nabisco	.38	.37	.42	.41	.39
Private label	.32	.33	.33	.33	.28
RMSE	—	.03	.07	.06	.06
Hit ratio (N = 486)	—	74%	48%	50%	61%
Loyals + Switchers					
Sunshine	.08	.08	.06	.07	.07
Keebler	.06	.07	.06	.05	.07
Nabisco	.56	.56	.56	.56	.55
Private label	.30	.29	.32	.32	.31
RMSE	—	.01	.03	.03	.02
Hit ratio (N = 1542)	—	78%	71%	71%	70%

[a]GL is Guadagni and Little (1983) model, KR is Krishnamurthi and Raj (1988) model, and AR is Allenby and Rossi (1991) model.
[b]RMSE is root means squared error.

Given the results on model validation, we now examine the semiparametric specification for its substantive marketing implications. Our discussion centers on the estimated brand choice elasticities and the distribution of heterogeneity in brand preferences. The own- and cross-elasticities[8] are reported in Table 3 for the semiparametric random effects specification.

Examining the own-price elasticities in Table 3, we find that Sunshine has the largest own-choice elasticity (−3.07), followed by Keebler (−2.55), private label (−1.33), and Nabisco (−1.02). We also see that the own-feature elasticities are all higher than the corresponding display elasticities. For example, the presence of a feature advertisement for Sunshine increases its choice probability by .64, whereas the effect for a special display is .37. Hence, a feature avertisement appears to be more effective in influencing the brand choice decision than a special display, though special displays were used more often than feature advertising for the product category.

Turning to the cross-choice elasticities, we see that when the unit price of Sunshine decreases by 1%, the probability of choosing Nabisco decreases by .29%, whereas when the price of Nabisco decreases by 1%, the probability of choosing Sunshine decreases by 2.08%. Likewise, when Sunshine is featured, the probability of choosing Nabisco decreases by .06, whereas when Nabisco is featured, the probability of choosing Sunshine drops by .31. This asymmetric pattern of competition between brands is generally consistent with the findings of Blattberg and Wisniewski (1989) that the competition between brands in different price tiers is asymmetric.

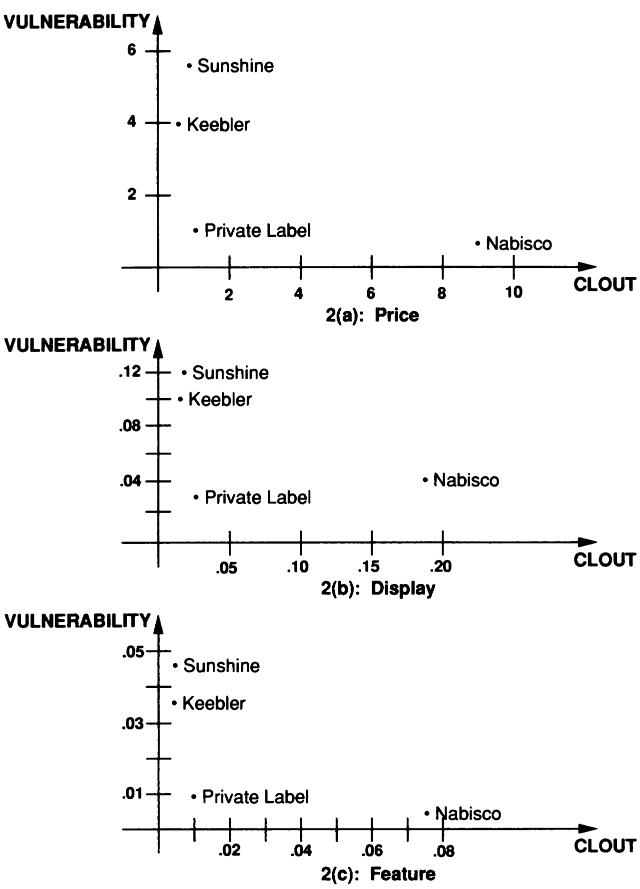
To summarize the pattern of competition between brands that is reflected by the estimated cross-elasticities, we plot in Figure 1 the "clout" and "vulnerability"

[8]For the dummy variables, feature and display, the elasticity represents the relative change in the probability due to a feature advertisement or special display.

Table 3
CHOICE ELASTICITIES FOR THE SEMIPARAMETRIC RANDOM EFFECTS MODEL

	Change in choice probability			
Change in	*Sunshine*	*Keebler*	*Nabisco*	*Private label*
Price elasticities				
Sunshine	−3.07	.53	.29	.19
Keebler	.46	−2.55	.23	.15
Nabisco	2.08	1.83	−1.02	.97
Private label	.60	.55	.45	−1.33
Feature elasticities				
Sunshine	.64	−.10	−.06	−.04
Keebler	−.10	.58	−.05	−.03
Nabisco	−.31	−.28	.17	−.17
Private label	−.11	−.10	−.08	.24
Display elasticities				
Sunshine	.37	−.06	−.03	−.02
Keebler	−.06	.33	−.03	−.02
Nabisco	−.19	−.17	.11	−.10
Private label	−.07	−.06	−.05	.14

Figure 1
CLOUT VERSUS VULNERABILITY



of each brand in terms of the different marketing mix variables (see Cooper and Nakanishi 1988 for a discussion of those two measures). We see from Figure 1 that Nabisco has the greatest clout and least vulnerability with respect to price, feature advertisement, and special display. Sunshine seems to be the brand that is most vulnerable to competitive actions and has the least clout.

Heterogeneity in Brand Preferences

For the semiparametric random effects specification, we find that three support points are sufficient to approximate the underlying probability distribution of intrinsic preferences for all four brands. The estimated support points and the associated probability masses are given in Table 4. To obtain greater insights about the nature of heterogeneity in preferences for the four brands, we plot the distributions for the four brands in Figure 2.

It is clear from Figure 2 that the probability distributions of intrinsic preferences are different for the various brands. This finding implies that the intrinsic preferences are both brand- and household-specific and hence it would not be appropriate to impose a common probability distribution for all brands. This is the advantage of using a flexible approach such as the semiparametric specification.

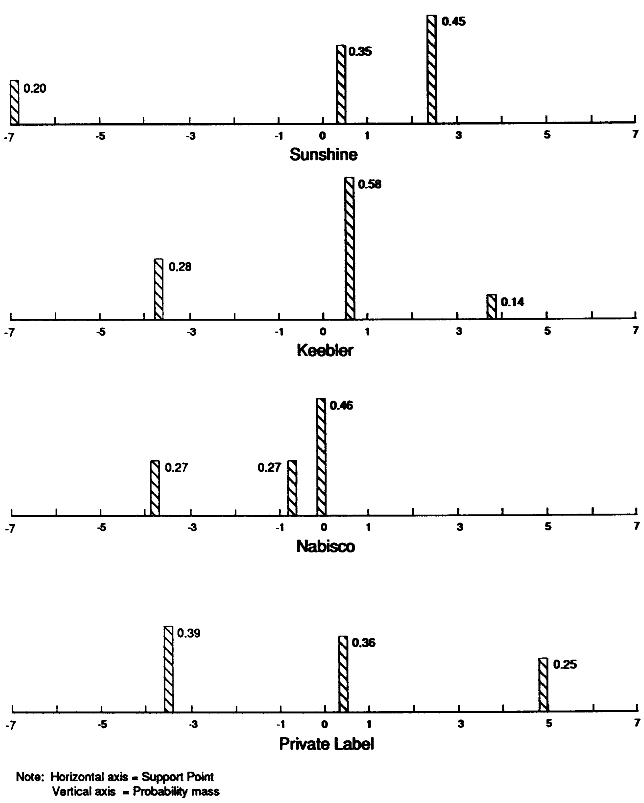
We also see from Figure 2 that though the distribution of preferences for Keebler is somewhat symmetric, the other three distributions seem to have more monotonic patterns. Further, examining the variances of the different distributions, we find that Sunshine has the greatest heterogeneity in intrinsic preferences (with a variance of 13.3) followed by private label (11.7), Keebler (6.1), and Nabisco (3.4). Hence, Nabisco has the least heterogeneity in preferences across households for these data. The mean intrinsic preference for each brand is given by the brand-specific constant (see equation 3). From Table 1, we see that Nabisco has the highest mean preference (4.527), followed by Keebler (.933), private label (0.0), and Sunshine (−.328).[9] It is interesting to note that the rank ordering of the mean intrinsic preferences is in-

[9]In estimating the brand-specific constants, the normalization was performed with respect to the private label brand.

Table 4
DISTRIBUTION OF INTRINSIC PREFERENCES

	First support point (ξ_1)		*Second support point* (ξ_2)		*Third support point* (ξ_3)	
Brand	*Location*	*Probability mass*	*Location*	*Probability mass*	*Location*	*Probability mass*
Sunshine	−7.13	.20	.81	.35	2.54	.45
Keebler	−3.59	.28	.78	.58	3.94	.14
Nabisco	−2.60	.27	−.44	.27	1.79	.46
Private label	−3.68	.39	.50	.36	5.02	.25

Figure 2
DISTRIBUTION OF INTRINSIC PREFERENCES



versely related to that of the variances of the distribution of heterogeneity for the four brands.

An interesting marketing issue is the interpretation of the distribution of intrinsic preferences across households. To address that issue, we first clarify the meaning of the estimated support points and probability masses associated with them. Consider, for example, the distribution of preferences for Keebler. The interpretation is that any randomly drawn household will have an intrinsic preference value of -3.59 with probability .28, a value of .78 with probability .58, and a value of 3.94 with probability .14. Alternatively, we can say that 28% of the households will have an intrinsic preference value of -3.59, 58% will have the value .78, and the remaining 14% will have the value of 3.94 for Keebler. Similar interpretations can be given for the other brands. Note that the preference points and their associated probability masses are different for the three brands. Hence, the semiparametric specification enables us to make inferences about the clustering of preferences of households (the location and the relative sizes of the clusters) for each brand.

SUMMARY AND DIRECTIONS FOR FUTURE RESEARCH

We show how heterogeneity in preferences across households can be incorporated in a logit model of brand choice using a random effects specification. We also compare the proposed approach with other methods used to account for heterogeneity in brand preferences.

An important empirical finding of our analysis is that the semiparametric random effects model seems to be the most appropriate specification to capture heterogeneity in brand preferences across households, in terms of both goodness-of-fit and predictive criteria. The distribution of intrinsic preferences that is captured by the semiparametric random effects specification is brand specific. This result indicates that in attempting to capture such heterogeneity, imposing a single probability distribution across all brands (e.g., Steckel and Vanhonacker 1988) would not be appropriate.

An important property of a discrete choice model is its ability to predict brand choice correctly in a holdout sample. Our analysis shows that the performance of the Guadagni and Little model decreases considerably in data with a high degree of switching behavior. In contrast, the predictive ability of the semiparametric random effect specification is robust to the nature of the brand choice behavior of households. Specifically, its performance in predicting brand choice is almost as good with data characterized by high repeat purchase as it is with data representing high brand switching behavior. This result is consistent with the theoretical findings of Dalal and Klein (1988) that a logit model with a "correct" specification of heterogeneity predicts aggregate brand choice probabilities well.

Our empirical analysis also indicates that for the brands considered in the saltine product category, the probability of choosing Sunshine is most affected when there is a change in the marketing variable of the other brands—that is, it is most vulnerable. Moreover, Sunshine also has the least clout among the various brands, implying that it draws less when it is on promotion. Nabisco seems to hold the dominant position in the product category, as measured by its clout and vulnerability.

The distribution of intrinsic preferences across households for the four brands indicates that Nabisco has the highest overall mean preference and the least heterogeneity, whereas Sunshine has the lowest mean brand preference but the highest level of heterogeneity. This finding together with those from the estimated elasticities suggest that Sunshine's performance in the marketplace is governed more by the nature and extent of the promotional activities undertaken than by its own consumer franchise.

Another source of heterogeneity is the variation across households in their response to the marketing mix variables. Though heterogeneity in the response coefficients is an important issue that warrants investigation, the difficulty lies in the estimation. The general procedure to handle such heterogeneity is to develop the model at the household level and then estimate the aggregate-level response coefficients by integrating with respect to the distribution of the parameters aross the households. That procedure would involve a multivariate distribution and hence the estimation would be extremely difficult (see Ben-Akiva and Lerman 1985).

To summarize, we attempt to address the important issue of accounting for heterogeneity in preferences in the analysis of panel data. We propose a new approach and compare it with other approaches in the literature. Such a comparative analysis should be helpful to researchers analyzing brand choice behavior in the context of frequently purchased package goods.

APPENDIX

For the parametric specification, the likelihood function in equation 13 does not have a closed-form expression for the integrals. To overcome that problem, we use Gaussian quadrature (Press et al. 1986) to approximate the integrals in the likelihood function. Because we assume that the distribution of heterogeneity for the N brands is independent, the integrals can be evaluated with respect to the distribution of one brand at a time. We first illustrate the technique for the case of a single brand.

The integrals we need to approximate are of the form

$$\text{(A1)} \qquad \int_0^\infty W(\xi) f(\xi) d\xi,$$

where:

$$W(\xi) = \left\{ \prod_{t=1}^{T_i} \left\{ \prod_{j=1}^{N} \left[\frac{\xi_j \exp(\mathbf{Z}_{ijt}\boldsymbol{\beta} + \theta_j)}{\xi_N + \sum_{k=1}^{N=1} \xi_k \exp(\mathbf{Z}_{ikt}\boldsymbol{\beta} + \theta_k)} \right]^{\delta_{ijt}} \right\} \right\}.$$

Using Gaussian quadrature, we can approximate the integral in A1 by the summation

$$\text{(A2)} \qquad \sum_{h=1}^{H} W(\xi_h)\lambda_h,$$

where the density function $f(\xi)$ is approximated by a finite number, H, of support points ξ_h with corresponding probability masses λ_h, $h = 1, 2, \ldots, H$. The function $W(\xi)$ is consequently evaluated only at each of these finite ξ_h values. Butler and Moffitt (1982) show that the quadrature yields good results even with a two-point approximation (i.e., $H = 2$). Therefore, it is necessary to obtain ξ_1, ξ_2 and the corresponding λ_1, λ_2 for this approximation.

When the random variable ξ in A1 is gamma distributed with parameters α and 1, Press et al. (1986) show that the support points ξ_h, $h = 1, 2, 3, \ldots, H$, are the roots of the Laguerre polynomial $l_H(\xi)$ of order H. The polynomial is given by the following recurrence relationship.

$$\text{(A3)} \qquad Hl_H(\xi) = (\alpha - 1 + 2H - \xi)l_{H-1}(\xi) - (\alpha + H - 1)l_{H-2}(\xi)$$

For $H = 2$, we have

$$\text{(A4)} \qquad 2l_2(\xi) = (\alpha + 3 - \xi)l_1(\xi) - (\alpha + 1)l_0(\xi).$$

From Table 22.4 of Abramowitz and Stegun (1972), we have

$$\text{(A5)} \qquad l_0(\xi) = 1; l_1(\xi) = (\alpha + 1 - \xi).$$

Substituting from A5 into A4 and simplifying, we have

$$\text{(A6)} \qquad l_2(\xi) = \frac{1}{2}\{\xi^2 - (2\alpha + 4)\xi + (\alpha^2 + 3\alpha + 2)\}.$$

The roots of the polynomial in A6 are given by

$$\text{(A7)} \qquad \xi_1 = (\alpha + 2) + \sqrt{\alpha + 2},$$
$$\xi_2 = (\alpha + 2) - \sqrt{\alpha + 2}.$$

To obtain the values of λ_h corresponding to the ξ_h's, we use the following result from Press et al. (1986).

$$\text{(A8)} \qquad \lambda_h = \frac{\phi_2(\xi_h)}{l_2'(\xi_h)}, \qquad h = 1, 2,$$

where l_2' is the derivative of l_2 with respect to ξ, and ϕ_2 follows the recurrence relationshp in A4 except that

$$\text{(A9)} \qquad \phi_0(\xi) = 0, \quad \phi_1(\xi) = l_1'.$$

Substituting from A9 into A4, we have

$$\text{(A10)} \qquad \phi_2(\xi) = \frac{1}{2}\{(\alpha + 3 - \xi)l_1'(\xi)\}.$$

Using the expression A10 in A8, we get

$$\text{(A11)} \qquad \lambda_1 = \frac{1}{2}\left[1 - \frac{1}{\sqrt{\alpha + 2}}\right]$$

and

$$\lambda_2 = \frac{1}{2}\left[1 + \frac{1}{\sqrt{\alpha + 2}}\right].$$

By repeatedly approximating all N integrals in the likelihood function in equation 13, we transform the integrals into summations that are amenable to use in standard maximum likelihood estimation routines.

REFERENCES

Abramowitz, Milton and Irene A. Stegun (1972), *Handbook of Mathematical Functions*. New York: Dover Publications, Inc.

Allenby, Greg (1990), "Hypothesis Testing With Scanner Data: The Advantage of Bayesian Methods," *Journal of Marketing Research*, 27 (November), 379–89.

——— and Peter Rossi (1990), "Quality Perceptions and Asymmetric Switching Between Brands," working paper, Ohio State University.

Bass, Frank M. (1974), "The Theory of Stochastic Preference and Brand Switching," *Journal of Marketing Research*, 11 (February), 1–20.

———, Abel P. Jeuland, and Gordon P. Wright (1976), "Equilibrium Stochastic Choice and Market Penetration Theories: Derivations and Comparisons," *Management Science*, 22 (June), 1051–63.

Ben-Akiva, Moshe and Steven R. Lerman (1985), *Discrete Choice Analysis*. Cambridge, MA: MIT Press.

Blattberg, Robert C. and Kenneth J. Wisniewski (1989), "Price-Induced Patterns of Competition," *Marketing Science*, 4 (Fall), 291–309.

Butler, J. S. and R. Moffitt (1982), "A Computationally Efficient Quadrature Procedure for the One-Factor Multinominal Probit Model," *Econometrica*, 50, 761–4.

Chamberlain, Gary (1980), "Analysis of Covariance With Qualitative Data," *Review of Economic Studies*, 47, 225–38.

Colombo, Richard and Jane Landwehr (1990), "A Non-Parametric Approach to Incorporating Heterogeneity Into the Logit Model," working paper, New York University.

Cooper, Lee and Masao Nakanishi (1986), *Market Share Analysis*. Boston: Kluwer Academic Press.

Dalal, Siddhartha R. and Roger W. Klein (1988), "A Flexible Class of Discrete Choice Models," *Marketing Science*, 7 (Summer), 232–51.

Gensch, Dennis H. and W. W. Recker (1979), "The Multinomial, Multiattribute Logit Choice Model," *Journal of Marketing Research*, 16 (February), 124–32.

Gonul, Fusun and Kannan Srinivasan (1990), "Modeling Unobserved Heterogeneity in Multinomial Logit Models: Methodological and Managerial Implications," working paper, Carnegie Mellon University, Pittsburgh.

Guadagni, Peter M. and John D. C. Little (1983), "A Logit Model of Brand Choice," *Marketing Science*, 2 (Summer), 203–38.

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

Hausman, Jerry, Bronwyn Hall, and Zvi Griliches (1984), "Econometric Models for Count Data With an Application to the Patents-R&D Relationship," *Econometrica*, 52, 909–38.

Heckman, James J. (1981), "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time–Discrete Data Stochastic Process," in *Structural Analysis of Discrete Data With Econometric Applications*, C. Manski and D. McFadden, eds. Cambridge, MA: MIT Press.

——— and Burton Singer (1984), "A Method of Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 52 (2), 271–320.

Horowitz, Joel (1980), "The Accuracy of the Multinomial Logit Model as an Approximation to the Multinomial Probit Model of Travel Demand," *Transportation Research*, 14B, 331–41.

Hsiao, Cheng (1986), *Analysis of Panel Data*. Cambridge, UK: Cambridge University Press.

Jain, Dipak and Naufel Vilcassim (1991), "Investigating Household Purchase Timing Decisions: A Conditional Hazard Function Approach," *Marketing Science*, 10 (Spring), 1–15.

Jones, J. Morgan and Jane T. Landwehr (1988), "Removing Heterogeneity Bias From Logit Model Estimation," *Marketing Science*, 7 (1), 41–59.

Krishnamurthi, Lakshman and S. P. Raj (1988), "A Model of Brand Choice and Purchase Quantity Price Sensitivities," *Marketing Science*, 7 (1), 1–20.

Kudpi, Vaman S. (1989), "Incorporating Heterogeneity in Brand Choice Models for Panel Data," unpublished doctoral dissertation, Northwestern University.

Lancaster, Tony (1979), "Econometric Models for the Duration of Unemployment," *Econometrica*, 47 (4), 939–56.

McFadden, Daniel (1974), "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers of Econometrics*, P. Zarembka, ed. New York: Academic Press, Inc., 105–42.

Press, W., B. Fleming, S. Tenkolsky, and W. Vetterling (1986), *Numerical Receipes*. Cambridge, UK: Cambridge University Press.

Srinivasan, V. and Thomas Kibarian (1990), "Purchase Event Feedback: Fact or Fiction?" paper presented at Marketing Science Conference (March 22–25), University of Illinois, Champaign-Urbana.

Steckel, Joel H. and Wilfred R. Vanhonacker (1988), "A Heterogeneous Conditional Logit Model of Choice," *Journal of Business and Economic Statistics*, 6 (3), 391–8.

Vilcassim, Naufel J. and Dipak C. Jain (1991), "Modeling Purchase-Timing and Brand-Switching Behavior Incorporating Explanatory Variables and Unobserved Heterogeneity," *Journal of Marketing Research*, 28 (February), 29–41.

Zufryden, Fred S. (1986), "Multibrand Transition Probabilities as a Function of Explanatory Variables: Estimation by a Least-Squares-Based Approach," *Journal of Marketing Research*, 23 (May), 177–83.

Reprint No. JMR284103