KAMEL JEDIDI, RAJEEV KOHLI, and WAYNE S. DESARBO*

The authors model product consideration as preceding choice in a segment-level conjoint model. They propose a latent-class tobit model to estimate cardinal, segment-level preference functions based on consumers' preference ratings for product concepts considered worth adding to consumers' self-explicated consideration sets. The probability with which the utility of a product profile exceeds an unobserved threshold corresponds to its consideration probability, which is assumed to be independent across product profiles and common to consumers in a segment. A market-share simulation compares the predictions of the proposed model with those obtained from an individual-level tobit model and from traditional ratings-based conjoint analysis. The authors also report simulations that assess the robustness of the proposed estimation procedure, which uses an E-M algorithm to obtain maximum likelihood parameter estimates.

Consideration Sets in Conjoint Analysis

Conjoint simulations are frequently used to predict the market share of new product concepts. Since the introduction of POSSE by Green, Carroll, and Goldberg (1981), many commercial products with simulation capabilities have been successfully introduced (e.g., Johnson 1987). A variety of extensions to the scope of the simulations have appeared in the marketing literature (e.g., Green and Krieger 1985; Kohli and Krishnamurti 1987; Kohli and Mahajan 1991; Kohli and Sukumar 1990; McBride and Zufryden 1985).

An alternative approach, first proposed by Louviere and Woodworth (1983), uses conjoint choice experiments to estimate preference functions at the aggregate or segment level (DeSarbo, Ramaswamy, and Cohen 1995). One advantage of this method is that by allowing a no-choice or current-choice option, it permits the modeling of consideration sets, in which an item is selected only if its utility exceeds a threshold (i.e., the utility of the no-choice alternative). As a consequence, the market share prediction for a new item can take into account both the consideration probabilities for distinct subsets of items and the choice probabilities of the items in a considered subset. In contrast, traditional conjoint simulations assume that a test product is always considered by each consumer. There is another advantage to this method: Because the number of consumers choosing an item depends on the considered set, choice-set experiments enable a new product to affect category penetration. Traditional conjoint allows for no such effect. This advantage of choice-set experiments is likely to be particularly important in markets in which a significantly improved new product can increase the number of category adopters.

*Kamel Jedidi is associate professor and Rajeev Kohli is professor, Graduate School of Business, Columbia University. Wayne S. DeSarbo is the Smeal Distinguished Chair of Marketing, Smeal College of Business, Pennsylvania State University.

The extensive use and enduring popularity of traditional conjoint analysis suggests that it may be useful to enhance it in a manner that allows for the inclusion of consideration sets in these models. We propose such a procedure, which, similar to choice-set experiments, models consideration as a function of product attributes. The proposed approach builds on the popular full-profile ratings-based conjoint analysis in two ways. First, we ask a consumer to specify which of the existing brands he or she considers in making a purchase. We assume that the selected brands are random draws from an unknown probability distribution that is identical across consumers within a market segment but can vary across segments. Thus, we do not require either the existence of a well-defined consideration set or the ability to enumerate all possible items that might be considered by a consumer. Instead, both the elicited set size and its composition are assumed to be random variables. Second, we sequentially present product profiles, while at each step asking the consumer to provide a preference rating only if the item is worth adding to the self-explicated consideration set. Thus, the data obtained are censored so that no ratings are obtained for product profiles not judged worth consideration by a consumer. We assume that (1) consideration judgments and preference ratings are simultaneously provided by the consumer, (2) each evaluation is independent but is anchored to the self-explicated consideration set, and (3) the preference evaluations are probabilistic and dependent on the composition of the described consideration set.

The data are used to estimate consideration and choice probabilities simultaneously in a linear-threshold model. However, in contrast to previous research that assumes the suitability of a linear consideration model (e.g., Roberts and Lattin 1991), we show that consideration based on a disjunctive rule or a conjunctive rule that is applied to all or a subset of attributes is a special case of a linear threshold model. Therefore, our approach is adequate for representing these two consideration processes.

Provided sufficient degrees of freedom are available, the model can be estimated at the individual level (Malhotra 1986). However, if, as is often the case, sufficient individual-level data are not available (e.g., if the subset of considered items is small, if the number of parameters is large compared to the number of profiles), we propose using a latent-class model that pools the data across consumers and simultaneously estimates consumer segments, segment-level part-worth functions, and segment-membership probabilities for consumers. In contrast to traditional conjoint share simulations (e.g., using a max-utility rule to simulate choices for all consumers), the present approach predicts market size (the proportion of consumers considering a product), conditions choice (and hence share) on consideration, and reflects estimation error in the predictions. Methodologically, the model generalizes the latent-class tobit model proposed by Jedidi, Ramaswamy, and DeSarbo (1993), which allows only one observation (consideration and product rating) per consumer. In the present conjoint setting, we permit multiple profile evaluations per respondent (repeated measurement) and use similarities in the within-subject response vectors to identify segments, part-worth functions, and segment membership probabilities for consumers.

First, we describe the proposed consideration model, its relationship to conjunctive and disjunctive processes, and the consequences of including consideration on market size and share predictions. Second, we present the new, segment-level model for analyzing the censored preferences. Third, we discuss the results from two simulations that use synthetic and commercial conjoint data. The first simulation assesses the robustness of the proposed procedure to violations of the model assumptions. The second simulation compares the segment and individual-level models with a traditional conjoint model that ignores the effect of consideration and performs share simulations by using the popular maximum utility rule.

THE CONSIDERATION MODEL, MARKET SIZE, AND MARKET SHARE

New products, especially when they offer substantial improvements over existing alternatives, can influence consideration by nonusers (e.g., cellular phones, satellite dishes) and frequency of brand consideration by category users (e.g., Parmalat milk, over-the-counter formulation of Pepcid, nondrowsy allergy and cold remedies, new age beverages such as Snapple). Process-tracing studies (e.g., Bettman 1979) suggest that consumers make consideration judgments using noncompensatory processes (Olshavsky and Acito 1980). In contrast, models for analyzing classification data (e.g., logit, discriminant analysis) assume a compensatory process, and have been used as approximate specifications in the consideration models proposed by Hauser and Wernerfelt (1990), Roberts and Lattin (1991), and DeSarbo and Jedidi (1995).

Cattin (1981) notes the suitability of a linear model for representing conjunctive preferences over discrete attributes. We show that both conjunctive and disjunctive processes can be represented as special cases of a linear threshold model.[1] The threshold of the linear model determines the rule used, and its parameters determine the acceptability of an attribute level.

To simplify the exposition, we begin by examining one brand (product profile) that is considered for purchase with probability p by a consumer. We assume that p is a function of the brand's utility u*, which in turn is a linear function of attributes, each of which is defined at a finite number of discrete levels; that is,

$$u^* = a + \sum_m b_m x_m + \epsilon, \tag{1}$$

where a is a constant, x_m is a suitable dummy variable identifying the attribute level, and ϵ is an error term.

Let w = 1 (0) denote that the brand is acceptable (unacceptable) to the consumer. We assume

$$w = \begin{cases} 1 \text{ if } u^* = a + \sum_m b_m x_m + \epsilon \geq T, \\ 0 \text{ otherwise;} \end{cases} \tag{2}$$

or, equivalently,

$$w = \begin{cases} 1 \text{ if } y^* = b_0 + \sum_m b_m x_m + \epsilon \geq 0, \\ 0 \text{ otherwise,} \end{cases} \tag{3}$$

where T is a utility threshold for the brand to be considered, $y^* = u^* - T$, and $b_0 = a - T$. Thus, the probability p that the brand is considered by the consumer is given by

$$p = P(w = 1) = P(y^* \geq 0) = P\left[\epsilon \geq -(b_0 + \sum_m b_m x_m)\right],$$

which increases with its utility $b_0 + \Sigma_m b_m x_m$ and depends on the density function of the error term, which we assume is normal with zero mean and variance σ^2. Thus, a brand has a consideration probability of $1 - \Phi[-(b_0 + \Sigma_m b_m x_m)/\sigma]$, where Φ is the cumulative normal density function. Note that ignoring consideration ($p = 1$) is equivalent to setting T (and thus b_0) to a large negative number compared to the parameters b_m.

We assume that the error is independent of the considered brand, and thus the brands have independent consideration probabilities. Noncompensatory consideration models such as the conjunctive and disjunctive rules also assume independent consideration of items, and traditional conjoint models assume that all items are considered. Nevertheless, the independence assumption implies the possibility of large consideration sets, which can be unreasonable in many cases. Although prespecified constraints on the set size can be imposed either at the estimation or the subsequent market

[1]The sufficiency of the linear model for representing lexicographic preferences is suggested in an example by Olshavsky and Acito (1980) and is readily established as follows: Let attribute k have n_k levels and let i = 1, 2, ..., n denote a decreasing lexicographic preference order for the $n = \Sigma_k n_k$ levels across attributes. Let level j_i of attribute k_i have rank i. Then the linear model $u = \Sigma_k b_{j_i k_i} x_{j_i k_i}$, where $x_{j_i k_i} = 1$ if level j_i of attribute k_i appears in the product and equals zero otherwise, represents a lexicographic preference order if its parameters satisfy the condition $b_{j_i k_i} > \Sigma_k \max_{\ell > i} \{b_{j_\ell k_\ell}\}$. It can be shown that the preferences u lie between an ordinal scale and an ordered semimetric (Coombs 1964).

simulation stage, this ad hoc approach does not reflect the effect of set size on the modeling of consideration.

A more reasonable approach is to assume that the consideration of an item depends on the size and composition of an existing consideration set. Moreover, it is also more realistic to assume that consumers have neither well-defined consideration sets nor the ability to enumerate all the items they might consider. Asked to identify considered items, a consumer probabilistically specifies a set $s = \{i \mid i = 1, 2, ...\}$. A reasonable assumption is that the more desirable the items in s (or the larger the size of s), the smaller the chance of a new item being considered by a consumer. Equivalently, the threshold T increases with the utility of the items in s and the size of s. Thus, we assume

$$T = c_0 + c_1 u_s + e_T,$$

where u_s is a measure of the desirability of s (i.e., $u_s = u(|s|, u_{is}, i = 1, 2, ...)$), c_0 is the threshold value when there are no previously considered items, and e_T is a random error term representing preference uncertainty (and in the subsequent segment-level model, within-segment variation in preferences). Equivalently,

$$(4) \qquad T = c_0 + c_1(\bar{u}_s + e_{u_s}) + e_T = c_0 + c_1\bar{u}_s + e,$$

where $\bar{u}_s$ is the true set-utility, e_{u_s} is an error term that captures uncertainty in u_s, and $e = c_1 e_{u_s} + e_T$. Thus, Equation 2 can be written as

$$(5) \qquad w = \begin{cases} 1 \text{ if } u^* = a + \sum_m b_m x_m + \epsilon \geq c_0 + c_1\bar{u}_s + e, \\ 0 \text{ otherwise;} \end{cases}$$

or

$$(6) \qquad w = \begin{cases} 1 \text{ if } y^* = b_0 + \sum_m b_m x_m + \epsilon' \geq 0, \\ 0 \text{ otherwise;} \end{cases}$$

where $b_0 = a - c_0 - c_1\bar{u}_s$ and $\epsilon' = \epsilon - e$. Note that $\bar{u}_s$ is absorbed in the additive constant and is not independently estimable. If sufficient degrees of freedom are available, the intercept b_0 can be estimated at the individual level and the remaining parameters at the segment level. Note that the individual-specific constant b_{0i} subsumes the effect of u_s (the utility the consumer obtains from an existing consideration set), which can depend on both the size and composition of the consideration set. Thus, a segment-level model that estimates individual-level intercepts reflects the possibility that the utility of an existing consideration set varies by consumer.

The linear threshold function of Equation 2 is sufficient for representing disjunctive and conjunctive processes. Let $\ell \geq 0$ denote the number of attribute levels acceptable to the consumer and let $m = 1, 2, ..., \ell$ identify these levels in Equation 2. Consider a special case of Equation 2 with the parameters constrained to have values $a = 0$ and

$$b_m = \begin{cases} 1 \text{ if } m = 1,2,...,1, \\ 0 \text{ otherwise;} \end{cases}$$

or, equivalently,

$$(2') \qquad w = \begin{cases} 1 \text{ if } u^* = \sum_{m=1}^{\ell} x_m + \epsilon \geq T, \\ 0 \text{ otherwise.} \end{cases}$$

If A denotes the total number of attributes, then $u^* = \sum_{m=1}^{\ell} x_m$ is an integer with a minimum value of zero if none of its attribute levels are acceptable ($\ell = 0$) and a maximum value of A if it is acceptable on all product attributes ($\ell = A$). Thus, a disjunctive rule that requires a considered alternative to have at least one acceptable attribute is equivalent to the condition

$$w = \begin{cases} 1 \text{ if } u^* \geq \ell, \\ 0 \text{ otherwise;} \end{cases}$$

and a conjunctive rule that requires all A attributes to be acceptable corresponds to the condition

$$w = \begin{cases} 1 \text{ if } u^* \geq A, \\ 0 \text{ otherwise.} \end{cases}$$

More generally, a consideration rule that requires an acceptable alternative to have at least k acceptable attributes is represented by a constrained linear threshold model, with the threshold taking a value of $T = k$. The addition of the error term in Equation 2' permits uncertainty or error in the threshold T, which is more accurate than a deterministic noncompensatory rule unfailingly used by a consumer.

Without the integer constraints, the model described by Equation 2' can be estimated using binary logit or probit. The constraints restrict the parameter space to $2^L A$ values, where L is the total number of levels across attributes, 2^L is the number of distinct beta vectors, and A is the number of possible threshold values. Thus, these two noncompensatory models correspond to a constrained specification of the linear threshold model.

Effect of Consideration on Market Size

A consumer asked to identify a consideration set selects a set c of items. Assume that this set is identified with an (unobserved) probability p_c. Presented with a new alternative j, the consumer considers it with probability p_j, given the set c (i.e., p_c is a conditional probability). Thus, the probability that at least one of the brands in $c \cup \{j\}$ is considered is

$$(7) \qquad p_c p_j + (1 - p_c)p_j + (1 - p_j)p_c = p_j(1 - p_c) + p_c.$$

Introducing brand j therefore increases the consideration probability of at least one brand by $d = p_j(1 - p_c)$ and increases the unconditional consideration probability by

$$\sum_{c \in C} p_j(1 - p_c),$$

where C is the set of all consideration sets c. In saturated markets (e.g., soft drinks), or in markets in which no short-term category-level alternatives exist (e.g., car batteries), the probability p_c for any consideration set c is likely to be close to one for most consumers, and a new brand is unlikely to affect category consideration (i.e., d is close to zero). However, if a category offers no particularly attractive alternative for a market segment (e.g., satellite dishes, electronic mail), and more attractive alternatives are available in a substitutable category (e.g., cable television, telephone, facsimile machine), then p_c is likely to be close to zero for consumers in these segments. An improved new product j (e.g., Sony's miniature satellite dish, Netscape) can increase category consideration, which tends to p_j as the consideration probability p_c of a subset c of existing brands tends to zero. Thus, the value of d is larger for better new products and is lower if the existing set of brands are attractive and/or a choice must be made.

Effect of Consideration on Market Share

Consideration can affect market share in two ways: by giving higher weight to choices in more probable consideration sets, and, when choice is not necessary, by eliminating all items from consideration. Specifically, let $s(j \mid c)$ denote the choice probability of item j from a given consideration set c. The unconditional choice probability of item j is

$$s_j = \sum_c p(c)\, s(j \mid c), \tag{8}$$

which weights the conditional choice probabilities $s(j \mid c)$ by the consideration probabilities $p(c)$. Let $p = \Sigma_c\, p(c)$ denote the probability of the consumer considering at least one item. Then, the unconditional share of choices for brand j is given by s_j/p.

Market share predictions obtained from traditional conjoint simulations differ in two ways from those obtained by the previous procedure. First, they assume $p = 1$, thereby requiring a consumer to consider at least one alternative (other values of p can be used, though the criteria and methods for selecting these value are unclear). Second, they permit only one consideration set—the set of all available items—and assume $p(c) = 0$ for all other consideration sets. Thus, the prediction error for traditional conjoint simulations due to any consideration effects is $(s_j/p) - s(j \mid c^*)$, where c* is the set comprising all alternatives. Note that the preceding development is easily generalized to account for segmentation and heterogeneity in consideration and choice probabilities.

ESTIMATING CONSIDERATION PROBABILITIES

Consider a consumer whose preference y* for a given product concept (or brand) is specified by Equation 3 or, equivalently, by Equation 6. Conditional on an existing consideration set, the proposed data-collection procedure requiring the customer to rate only considered brands provides a preference rating y = y* if the product profile is considered acceptable and a censored value y = 0 if the brand is not considered. Assuming that the untruncated preference y* has a normal distribution, the unconditional density function for the censored preference y is given by (see Amemiya 1984)

$$h(y) = \frac{1}{\sigma}\phi(z)^w\, \Phi(z)^{1-w}, \tag{9}$$

where ϕ and Φ denote the density and cumulative distribution functions of the standardized normal variate

$$z = \frac{y - (b_0 + \Sigma_m b_m x_m)}{\sigma}.$$

Let $y = \{y_j, j = 1, 2, ..., J\}$ denote the censored preferences a consumer provides for a set of J product profiles. Because the preference ratings are assumed to be independent, the joint density function for the censored ratings is given by $\Pi_j\, h_j(y_j)$, where the subscript j identifies the product profiles.

The parameters b_m of a consumer's preference model can be obtained by maximizing the likelihood function

$$\mathcal{L}_1 = \Pi_j h_j(y_j) = \Pi_j \frac{1}{\sigma}\, \phi(z_j)^{w_j}\, \Phi(z_j)^{1-w_j}, \tag{10}$$

where $w_j = 1$ (0) if the consumer considers (does not consider) brand j. Alternatively, if preference homogeneity is assumed across a sample of I consumers, the parameters of an aggregate preference model can be estimated by maximizing the likelihood function

$$\mathcal{L}_2 = \Pi_{i,j}\, h_j(y_{ij}) = \Pi_{i,j} \frac{1}{\sigma}\, \phi(z_{ij})^{w_{ij}}\, \Phi(z_{ij})^{1-w_{ij}}, \tag{11}$$

where y_{ij} is consumer i's censored preference rating for product profile j, h_j is the density function for profile j and is common across consumers, and $w_{ij} = 1$ (0) if product profile j is acceptable (not acceptable) to consumer i. Furthermore, if individual membership in K prespecified segments is available, the estimates of the parameters for the segment preference functions can be obtained by maximizing the likelihood function

$$\mathcal{L}_3 = \Pi_{i,j,k}\, [h_{jk}(y_{ijk})]^{d_{ik}} = \Pi_{i,j,k}\left[\frac{1}{\sigma_k}\, \phi(z_{ijk})^{w_{ijk}}\, \Phi(z_{ijk})^{1-w_{ijk}}\right]^{d_{ik}}, \tag{12}$$

where y_{ijk} is the censored preference rating for product profile j given by consumer i in segment k, h_{jk} is the density function for profile j and segment k, $w_{ijk} = 1$ (0) if product profile j is acceptable (not acceptable) to consumer i in segment k, $d_{ik} = 1$ (0) if consumer i belongs (does not belong) to segment k, and[2]

$$z_{ijk} = \frac{y_{ijk} - (b_{0k} + \Sigma_m b_{mk} x_{jm})}{\sigma_k}. \tag{13}$$

Finally, in the most general case, in which segment membership is unknown, assume that all consumers initially have the same probability λ_k of membership in segment k, k = 1, 2, ..., K; that is, $P(d_{ik} = k) = \lambda_k$ for all i. Because the density of y_i conditional on membership in segment k is $\Pi_j\, h_{jk}(y_{ij})$, its unconditional distribution is a finite mixture of such densities. That is,

$$h_m(y_i) = \sum_k \lambda_k\, \Pi_j\, h_{jk}(y_{ij}). \tag{14}$$

The estimation of both the segment partworths b_{mk} and the "mixing proportions" λ_k can be achieved by maximizing the likelihood function,

$$\mathcal{L}_4 = \Pi_i\, h_m(y_i). \tag{15}$$

Because the number of segments is generally unknown, $\mathcal{L}_4$ must be maximized for various values of K. The optimal number of segments is selected on the basis of the minimum Consistent Akaike Information Criterion (CAIC)

$$\text{CAIC} = -2\ln \mathcal{L}_4(K) + N_K(1 + \ln I), \tag{16}$$

where $N_K = JK + 2K - 1$ is the number of effective parameters, $-2\ln \mathcal{L}_4(K)$ is the usual χ^2 measure for the likelihood

[2]If sufficient degrees of freedom are available, the intercept term can be estimated at the individual level (i.e., b_{0i}) by introducing an additional I – 1 consumer-specific dummy variables in the utility equation.

function, and $N_K(1 + \ln I)$ is a penalty term that accounts for the improvement in fit obtained by increasing the number of parameters and the number of data points in the model (see, e.g., Bozdogan 1987).

With $h_{jk}(y_{ij})$ and λ_k, Bayes rule can be used to estimate the posterior probability P_{ik} that consumer i belongs to segment k, where (see Jedidi, Ramaswamy, and DeSarbo 1993)

$$P_{ik} = E(d_{ik} \mid y_i) = \frac{\lambda_k \Pi_j h_{jk}(y_{ij})}{\Sigma_\ell \lambda_\ell \Pi_j h_{j\ell}(y_{ij})}. \quad (17)$$

These posterior probabilities, when evaluated at the maximum likelihood estimates $\hat{\lambda}_k$, $\hat{b}_{mk}$, and $\hat{\sigma}_k$, k = 1, 2, ... K, m = 1, 2, ... M, represent a fuzzy classification (clustering) of I consumers into K latent segments. The adequacy of the estimated membership probabilities is assessed through an entropy measure

$$E = 1 - \frac{\Sigma_{i,k} - P_{ik} \ln P_{ik}}{I \ln K}, \quad (18)$$

which has a value of one when all membership probabilities are either zero or one and equals zero when all membership probabilities are equal to 1/K.

A brief overview of the E-M algorithm developed to estimate the segment memberships and partworth simultaneously is subsequently described.[3] This algorithm is general because it can be used to perform individual- ($\mathcal{L}_1$), segment- ($\mathcal{L}_3$ and $\mathcal{L}_4$), and aggregate-level ($\mathcal{L}_2$) conjoint analyses when preference ratings are truncated or untruncated and when segment memberships are known or unknown.

The E-M Algorithm

The E-M algorithm, which is suited for models with censored and missing data, is a general iterative method for obtaining maximum likelihood estimates (Dempster, Laird, and Rubin 1977; Goodman 1974; Grover and Srinivasan 1987; Malhotra 1986; Zenor and Srivastava 1993). The E-M method cycles between an expectation phase in which the missing data (segment membership and censored observations) are replaced by their expected values, given provisional estimates for b_{mk}, σ_k^2, and λ_k, and a maximization phase in which the "full" data are then used to obtain new estimates for the model parameters. This iterative process continues until no further improvement in the likelihood function is possible.

Schematically, the proposed E-M algorithm can be described as follows:

Initialization step. Randomly generate initial estimates of the segment partworth b_{mk}, the variance σ_k^2 of the preference functions, and the mixing proportions λ_k, k = 1, 2, ..., K. Using Equation 17, estimate $P_{ik} = E(d_{ik} \mid y_i)$ for all consumers and segments. Use the following equation to estimate the expected value $E(y^*_{ij} \mid y_{ij} = 0)$ for product profile j, which is censored (not considered) by consumer i:

$$E(y^*_{ij} \mid y_{ij} = 0) = \sum_k P_{ik} \left\{ \sum_m b_{mk} x_{jm} - \sigma_k \frac{\phi(z_{jk})}{1 - \Phi(z_{jk})} \right\}, \quad (19)$$

where $z_{jk} = \Sigma_m b_{mk} x_{jm} / \sigma_k$.

Let

$$\hat{y}_{ij} = \begin{cases} y_{ij}, & \text{if } y_{ij} > 0; \\ E(y^*_{ij} \mid y_{ij} = 0) & \text{if } y_{ij} = 0. \end{cases} \quad (20)$$

Thus, $\hat{y}_{ij}$ equals the actual (expected) value of profile j's rating if consumer i considers (does not consider) profile j acceptable in the profile evaluation task.

Recursion step. Use the updated membership probabilities and the updated values of the y_{ij} to reestimate the following parameters:

$$\beta_k = \frac{(X'X)^{-1} X' \Sigma_i P_{ik} \hat{y}_i}{\Sigma_i P_{ik}}; \quad (21)$$

$$\sigma_k^2 = \sum_{ij} \frac{P_{ik}[(\hat{y}_{ij} - \Sigma_m b_{mk} x_{jm})^2 + \text{var}(\hat{y}_{ij})]}{\Sigma_{i,j} P_{ik}}; \quad (22)$$

and

$$\lambda_k = \sum_i \frac{P_{ik}}{I}, \quad (23)$$

where $\beta_k = \{b_{mk}, m = 0, 1, ..., M\}$, $X = \{x_{jm}, j = 1, 2, ..., J; m = 0, 1, ..., M\}$, and

$$\text{var}(\hat{y}_{ij}) = \begin{cases} \sum_k P_{ik} \left\{ \sigma_k^2 + \left[\sum_m b_{mk} x_{jm} \sigma_k \frac{\phi(z_{jk})}{1 - \Phi(z_{jk})} \right] - \left[\sigma_k \frac{\phi(z_{jk})}{1 - \Phi(z_{jk})} \right]^2 \right\} & \text{if } y_{ij} = 0; \\ 0 & \text{if } y_{ij} = 0 \end{cases} \quad (24)$$

Reestimate $E(d_{ik}) = P_{ik}$ and $\hat{y}_{ij}$ using the new estimates of b_{mk}, σ_k, and λ_k.

Termination step. Stop when the increase in the likelihood function value $\mathcal{L}_4$ in Equation 15 is below a threshold, or if the number of iterations reach a preset maximum.

As other researchers have noted (e.g., Grover and Srinivasan 1987; Zenor and Srivastava 1993), an E-M algorithm can converge to local optima and therefore is typically implemented using multiple starting points. All simulations reported here use ten random starting points. The solution maximizing the likelihood function is retained as the final solution. Also, though the E-M algorithm can in general converge slowly, it performed reasonably well here, with the maximum time for obtaining a final solution being 60 seconds.

SIMULATIONS

Two simulations were performed: The first assessed the ability of the segment-level (latent class) model to recover preference functions and segment membership, and the second tested the relative performance of the segment- and individual-level tobit models, as well as the traditional conjoint model, to predict the market performance of a new product.

[3] The details of the derivations are available in a working paper.

Objectives and Design of the First Simulation

An estimation procedure should at least be able to accurately recover segment membership, consideration probabilities, and preference ratings from sample data. Furthermore, the procedure should be robust in recovering this information from noisy data, data in which the subset of considered items is small, and data that violate the assumption of normal errors. Our objective of this simulation is therefore to perform a "stress test" of the proposed estimation procedure, assuming the proposed model accurately reflects consideration and consumer preferences. Thus, using the following six factors and two replications per cell, we generated 972 problems according to a 2×3^5 factorial design:

1. Misspecification of the error distribution (normal, uniform, exponential),
2. Alternative levels of error variance (10%, 20%, 30% of the true variance),
3. Changes in proportions of unacceptable (censored) products (20%, 40%, 60%),
4. Alternative sample sizes (100, 300 consumers),
5. Variations in the numbers of segments (2, 4, 6) with different average preferences, and
6. Differences in the number of design (dummy) variables (4, 8, 12).

The "true" partworth for segment k was identically and independently generated from a uniform distribution over the range [–k, k]. For each simulation run, the true segment ratings were computed for 16 orthogonal product profiles. The profile ratings for consumers in a segment were generated by adding random observations from the appropriate (uniform, exponential, or normal) distribution; its variance, which was determined by the simulation trial, was 10%, 20%, or 30% of the variance in the true segment ratings of the 16 profiles (Srinivasan 1975). The appropriate proportion of product profiles (20%, 40%, or 60%) with the lowest observed ratings were censored for each consumer. These data were used to estimate the segment-level model.

Results of the First Simulation

The correlation between the true and estimated partworth has a mean value of .979, and is significantly reduced by data censoring ($p \leq .05$); its mean value decreased from .983 with 20% censoring (unacceptable) to .975 with 60% censoring. The correlation between the actual (0 or 1) and estimated posterior probabilities of segment membership has an average value of .953 across trials and is significantly affected at the 5% level by the number of segments, the number of design variables, and the error distribution. The most important among these is the error distribution; the lowest mean correlation was .905 for the exponential distribution. Variations in the proportion of unacceptable profiles (data censoring), error magnitude, and sample size have no significant effects on the correlation.

Finally, 89.3% of the product profiles were correctly classified into acceptable and unacceptable categories, with only data censoring having a significant effect ($p \leq .05$) on the correct classification percentage, which drops to 88.1% for 60% truncation. Overall, the performance of the proposed procedure is reasonable on this classification measure. Thus, this initial test of the estimation procedure suggests that it performs well in recovering the preference functions, predicting consideration, and identifying the number and membership of market segments.

Objectives and Design of the Second Simulation

Adequate recovery of parameters notwithstanding, it is not clear if the predictions of a traditional conjoint model differ substantially from those of the proposed model. If the traditional conjoint model is sufficiently robust to yield share predictions close to those obtained from the proposed model, then the additional complexity of the latter may not be justifiable from a predictive standpoint, even if the proposed model provides a more faithful representation of consumer choice processes. Also, the segment-level tobit gains observations by pooling across respondents, whereas the individual-level tobit is better at capturing idiosyncratic preferences. It is therefore important to compare the two models in terms of their ability to predict market size and market share.

A second simulation was performed to address the previous issues. Specifically, our objectives are to (1) compare the standard conjoint model with the individual- and segment-level tobit models in terms of market share predictions and (2) assess the ability of the two (individual- and segment-level) tobit models in capturing the market penetration achieved by all products on the market.

The number of profile evaluations (16, 32) and the proportion of unacceptable products (20%, 40%, 60%) were varied across the simulation runs. Censored preference ratings for 32 product profiles described over six attributes (four with two levels, and one each with three and four levels) were collected for a commercial conjoint study. These data were analyzed for 211 respondents through the segment-level tobit model.[4]

Four segments comprising 36%, 18.5%, 28%, and 17.5% of the consumers were identified through CAIC. The estimated values of the consideration probabilities (which across segments have average values ranging from .20 to .81), the segment part worth (b_{mk}), and the error variances (s_k^2) of the segment preference functions were treated as the simulation parameters. These parameters were used to generate segment k's true rating for product profile j (y^*_{jk}); the selected profiles were specified by a simulation trial. A random observation from $N(0, s_k^2)$ was added to segment k's true preference score to generate consumer i's rating y^*_{ijk} for product profile j. These uncensored profile ratings were used as the dependent variables in a least-squares regression to estimate individual-level preference functions for the standard conjoint model. Ten replications were used per treatment condition, which resulted in $2 \times 3 \times 10 = 60$ simulated data sets. The individual- and segment-level tobit models were estimated after the appropriate percentage of consumer responses were censored.

Nine choice sets, each comprising three items, were generated to examine the effects of various combinations of consideration probabilities on the share predictions of the models. The mean (across segments) values of the consideration probabilities for the items in each choice set S_i are as follows:

[4]The details of this empirical application are available in a working paper.

$S_1 = \{.20, .24, .27\}$
$S_2 = \{.20, .23, .47\}$
$S_3 = \{.20, .47, .47\}$
$S_4 = \{.47, .47, .51\}$
$S_5 = \{.20, .24, .81\}$
$S_6 = \{.20, .47, .81\}$
$S_7 = \{.47, .47, .81\}$
$S_8 = \{.47, .81, .81\}$
$S_9 = \{.80, .81, .81\}$

For each choice set, the two tobit models were used to simulate the segment penetration across products, and all three models were used to simulate the market shares.

Market segment penetration for a choice set S_i is computed as

$$p_k(S_i) = 1 - \prod_{j \in S_i} (1 - p_{jk}),$$

where p_{jk} is the consideration probability of product profile j in segment k. Share of a product within a segment is obtained by first using Equation 8 to compute the within-segment unconditional choice probability and then normalizing this quantity by $p_k(S_i)$. That is,

$$s_{jk}(S_i) = \frac{\sum_{c \in C_i} p_k(c) s_k(j \mid c)}{p_k(S_i)},$$

where C_i is the set of all possible consideration sets in S_i and $s_k(j \mid c)$ is the conditional probability that product profile j has the maximum utility in consideration set $c \in C_i$ for segment k (i.e., $P[u_j > u_l, l \neq j]$). The market-level penetration and share estimates are obtained by weighting the corresponding segment-level values by the mixing proportions λ_k. Note that though presented at the segment level, this approach for computing penetration and shares is also applicable for individual as well as aggregate analyses. Furthermore, this approach is preferable to either ignoring error (as in the max-utility rule) or using probabilistic rules that are independent of the errors (e.g., Bradley-Terry-Luce (BTL) and logit computed using the partworth estimates).[5] Finally, recall that traditional conjoint analysis assumes $p_k(S_i) = 1$ and only uses the set of all items $c^* = S_i$ as the consideration set. Thus, using our approach to predict share in traditional, ratings-based conjoint amounts to computing the probability that a product profile has the maximum utility for each consumer (segment) and then averaging (weighting) these quantities to obtain market share.

Results of the Second Simulation

The true market penetrations for S_1 through S_9 are shown in the first panel of Table 1 and vary from 55.6% for S_1 to 99.3% for S_9. The values of the normalized root mean square deviation between the true and estimated market penetration across the treatments for the individual- and segment-level tobit models also are shown in Table 1. On average, the root mean square deviation has a value of 8.85% for the individual-level tobit model and 4.35% for the segment-level tobit model.

[5]It is easy to show that conjoint data are unsuitable for both BTL, which requires ratio utilities, and logit, which permits only the addition of a constant term to the utilities.

Table 1
ACCURACY OF PREDICTED MARKET PENETRATION: ROOT MEAN SQUARES DEVIATION*

Panel 1			
Choice Set	*True Market Penetration*	*Individual-Level Tobit*	*Segment-Level Tobit*
S1	55.6%	4.7%	1.8%
S2	67.3	6.4	3.1
S3	77.5	6.2	3.2
S4	86.2	9.9	5.2
S5	88.4	8.4	4.2
S6	91.9	7.5	4.1
S7	94.7	10.7	5.4
S8	98.0	12.6	6.0
S9	99.3	13.2	6.5
Panel 2			
Number of Product Profiles		*Individual-Level Tobit*	*Segment-Level Tobit*
32		8.0%	3.6%
16		9.7	5.1
Panel 3			
Percentage of Unacceptable Product Profiles		*Individual-Level Tobit*	*Segment-Level Tobit*
20%		8.9%	3.6%
40		9.4	4.2
60		8.3	5.3

*The cell entries are the percentage values of the root mean squares deviation between the true and estimated market sizes.

An analysis of variance performed on the arc-sin transformed values of the normalized root mean squares confirms that the performance of both tobit procedures deteriorates with reductions in the number of product profiles evaluated, that the accuracy of the predictions varies across choice sets, and that the segment-level tobit model better approximates the actual market size (all F-tests were significant at the 1% level). Note, however, that both models perform reasonably well, that the market size reduction can be large, and that either model is a substantial improvement over current simulations that ignore market size effects.

The share prediction for the two tobit models were obtained by computing the average value of the individual or segment choice probabilities, with the latter weighting the segment choice probabilities by the mixing proportions λ_k. Across treatment conditions, the root mean squared deviation between the predicted and true market shares is 32.45% for the standard conjoint model, 12% for the individual-level tobit model, and 5% for the segment-level tobit model.

In Table 2, we show the variation in the market share predictions across the treatments for the three models. Both tobit models perform significantly better than the deterministic simulation that ignores the effect of consideration, which is a conclusion supported by an analysis of variance of the arc-sin transformed values of the normalized root mean squares ($p \leq .01$). A Tukey pairwise comparison of the arc-sin transformed root mean squares also indicates that the segment-level tobit model provides closer market share estimates than the individual-level tobit model ($p \leq .05$), especially if the data are highly censored. An analysis of the (z-transformed) correlation between the true and estimated

Table 2

TRUE AND ESTIMATED SHARE: ROOT MEAN SQUARES (RMS) AND CORRELATION COEFFICIENT (CORR)

Panel 1

	True Market Share of Brand			*Traditional Conjoint*		*Individual-Level Tobit*		*Segment-Level Tobit*	
Choice Set	*1*	*2*	*3*	*RMS**	*CORR*	*RMS*	*CORR*	*RMS*	*CORR*
S1	.27	.28	.44	30.5%	.76	11%	.95	6%	.98
S2	.11	.13	.76	41.5	.67	15	.93	4	1**
S3	.04	.05	.91	29.5	.72	10	.96	4	1
S4	.45	.43	.12	36.0	.76	9	.97	6	.99
S5	.30	.30	.40	25.5	.88	10	.98	6	.98
S6	.11	.10	.79	26.5	.79	11	.98	5	.99
S7	.50	.48	.02	32.5	.73	12	.96	5	1
S8	.47	.44	.09	42.5	.49	12	.95	6	.99
S9	.31	.28	.41	30.5	.74	15	.76	8	.98

Panel 2

	Traditional Conjoint		*Individual-Level Tobit*		*Segment-Level Tobit*	
Number of Product Profiles	*RMS*	*CORR*	*RMS*	*CORR*	*RMS*	*CORR*
32	31.9%	.73	11%	.93	6%	.99
16	33	.72	13	.90	4	.99

Panel 3

	Individual-Level Tobit		*Segment-Level Tobit*	
Percentage of Unacceptable Product Profiles	*RMS*	*CORR*	*RMS*	*CORR*
20%	9%	.97	3%	.996
40%	11	.96	5	.992
60%	14	.89	8	.979

*The cell entries are the values of the percentage of the root mean squares deviation between the true and estimated market shares.

**All values are rounded off to the second decimal place.

market shares for the three models, which are also reported in Table 2, supports the previous conclusions. Across treatments, the correlation between the true and estimated market shares is .73 for the deterministic conjoint simulations, .94 for the individual-level tobit, and .99 for the segment-level model.

In summary, the market shares predicted by the traditional conjoint and tobit models are substantially different. Which model is closer to real consumer decision processes is an open question, but the results do not support the use of the simpler traditional model on the basis of predictive similarity alone. Between the individual- and segment-level tobit models, however, the latter appears to provide better share predictions and marginally better market penetration predictions—assuming the accuracy of the proposed consideration model. Overall, the segment-level model appears to outperform the individual-level tobit model, and both produce share predictions substantially different from the traditional conjoint model.

CONCLUSION

Our analysis suggests that, as in conjoint choice experiments, including the effect of consideration can improve the predictive performance of the conjoint choice simulators used for predicting the effect of product introductions and changes on market share and can provide an estimate of the potential changes in market penetration that occur as a consequence of brand introductions, deletions, or reformulations. Both the segment- and individual-level tobit models perform better than the commonly used max-utility model, which ignores consideration as well as the error component in the preferences.

The proposed segment-level tobit model performs better than the individual-level tobit model in large part because it pools information across customers with similar preferences. The limited degrees of freedom for the individual-tobit model can result in larger standard errors for the parameters than those obtained with the segment-level model. On the other hand, the segmentation itself is important, because aggregate estimation can significantly decrease the standard errors due to the pooling of heterogeneous preferences. In this sense, the segment-level tobit model is a useful extension of the existing estimation procedures because it reflects the advantages of both segment-level conjoint models and individual-level models that take into account the effect of consideration.

An additional benefit of the segment-level model is that as it pools observations across consumers with similar preferences, it offers the potential for reducing the amount of data collected from a respondent. This feature is likely to be increasingly valuable as firms begin to implement conjoint studies with large numbers of attributes. Both respondent effort and data-collection costs can be potentially reduced by using block designs to identify the profiles evaluated by a consumer, with the master design itself permitting estimates of main and interaction effects at the segment level.

There are at least six important areas that need further development. First, further research should develop and test alternative consideration set measures that do not assume independence of the evaluations. Second, the more technical issue is the development of associated predictive models that reflect the interdependence of items in consideration sets. Third, our model assumes the same utility function for consideration and preference. An alternative specification is to model the two decisions separately, in which a random utility approach is used to model consideration (binary choice) and a linear multiattribute model is used to represent preferences conditional on consideration (see Maddala 1983, p. 224). Thus, it is important to extend this two-stage model by using a latent class framework to estimate segment-level parameters. It also would be interesting to examine the robustness of our model in selecting the correct number of segments and in predicting market share as compared to a two-stage model. Fourth, as was suggested by our formulation of conjunctive and disjunctive processes, it may be useful to develop explicit tests for alternative consideration processes in future models. Fifth, though we have focused largely on the problem of incorporating consideration sets into full-profile rating-based conjoint analysis, it would be interesting and useful to compare our model to choice-based conjoint models (e.g., DeSarbo, Ramaswamy, and Cohen 1995; Louviere and Woodworth 1983), especially because these models permit the specification of a "choose-none" alternative. Sixth, we do not assess if the proposed model makes better share predictions than the standard conjoint model in a real market situation. A strong test would be to compare the predictions of the two models with actual choice. A weaker test would be to compare them on the basis of the prediction of revealed brand choices.

REFERENCES

Amemiya, Takeshi (1984), "Tobit Models: A Survey," *Journal of Econometrics*, 24, 3–61.

Bettman, James R. (1979), *An Information Processing Theory of Consumer Choice*. Reading, MA: Addison-Wesley.

Bozdogan, Hamparsam (1987), "Model Selection and Akaike's Information Criterion (AIC): The General Theory and Its Analytical Extension," *Psychometrika*, 52, 345–70.

Cattin, Philippe (1981), "On the Use of Verbal Protocols in Conjoint Analysis Studies," 12 (January), 715 – 717.

Coombs, C. H. (1964), *A Theory of Data*. New York: John Wiley & Sons.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977), "Maximum Likelihood From Incomplete Data via the E. M. Algorithm," *Journal of the Royal Statistical Society*, B39, 1–38.

DeSarbo, Wayne S. and Kamel Jedidi (1995), "The Spatial Representation of Consideration Sets," *Marketing Science*, 14, 326–42.

———, Venkatram Ramaswamy, and Steven Cohen (1995), "Marketing Segmentation With Choice-Based Conjoint Analysis," *Marketing Letters*, 6, 137–48.

Goodman, Leo A. (1974), "Exploratory Latent Structure Analysis Using Both Identifiable and Unidentifiable Models," *Biometrika*, 61 (2), 215–31.

Green, Paul E., J. Douglas Carroll, and Stephen M. Goldberg (1981), "A General Approach to Product Design Optimization via Conjoint Analysis," *Journal of Marketing*, 45 (3), 17–37.

——— and Abba M. Krieger (1985), "Models and Heuristics for Product Line Selection," *Marketing Science*, 4, 1–19.

Hauser, John R. and Berger Wernerfelt (1990), "An Evaluation Cost Model of Evoked Sets," *Journal of Consumer Research*, 16, 393–408.

Grover, Rajiv and V. Srinivasan (1987), "A Simultaneous Approach to Market Segmentation and Market Structuring," *Journal of Marketing Research*, 24 (May), 139–53.

Jedidi, Kamel, Venkatram Ramaswamy, and Wayne S. DeSarbo (1993), "A Maximum Likelihood Method for Latent Class Regression Involving a Censored Dependent Variable," *Psychometrika*, 58, 375–94.

Johnson, Richard D. (1987), "Adaptive Conjoint Analysis Software," in *Sawtooth Software*, Sequim, WA.

Kohli, Rajeev and Vijay Mahajan (1991), "A Reservation-Price Model for Optimal Pricing of Multi-Attribute Products in Conjoint Analysis," *Journal of Marketing Research*, 28 (August), 347–54.

——— and Ramesh Krishnamurti (1987), "A Heuristic Approach to Product Design," *Management Science*, 36, 311–22.

——— and R. Sukumar (1990), "Heuristics for Product-Line Design Using Conjoint Analysis," *Management Science*, 36, 1464–78.

Louviere, Jordan J. and George Woodworth (1983), "Design and Analysis of Simulated Consumer Choice on Allocation Experiments: An Approach Based on Aggregate Data," *Journal of Marketing Research*, 20 (November), 350–67.

Maddala, G. S. (1983), *Limited Dependent and Qualitative Variables in Econometrics*. New York: Cambridge University Press.

Malhotra, Naresh K. (1986), "An Approach to the Measurement of Consumer Preferences Using Limited Information," *Journal of Marketing Research*, 23 (February), 33–40.

McBride, R. D. and Fred S. Zufryden (1988), "An Integer Programming Approach to Optimal Product-Line Selection," *Marketing Science*, 7 (2), 126–40.

Olshavsky, Richard W. and Franklin Acito (1980), "An Information Processing Probe Into Conjoint Analysis," *Decision Sciences*, 11 (July), 451–70.

Roberts, John H. and James M. Lattin (1991), "Development and Testing of a Model of Consideration Set Composition," *Journal of Marketing Research*, 28 (November), 429–40.

Srinivasan, V. (1975), "Linear Programming Computational Procedures for Ordinal Regression," *Journal for the Association for Computing Machinery*, 23, 475–87.

Zenor, Michael J. and Rajendra K. Srivastava (1993), "Inferring Market Structure With Aggregate Data: A Latent Segment Approach," *Journal of Marketing Research*, 30 (August), 369–79.