JAIHAK CHUNG and VITHALA R. RAO*

The attribute-based approach to study customer choices cannot deal with bundles of heterogeneous components, which are usually drawn from different product categories. The authors develop the comparability-based balance model, which is a unified framework for modeling bundle choices. The model can be employed for any bundle, regardless of the heterogeneity of bundle components, under a pure bundling strategy. The conjoint model and balance model are special cases of the general model. The authors use mixture distributions in a hierarchical Bayesian framework to incorporate consumer heterogeneity in a more flexible manner than the extant approaches such as latent-class and random-coefficient models. The empirical tests of the model show that most attribute effects are significant and consistent with the model's predictions. The model is superior to those that do not consider the issues of comparability of attributes, latent-class structure, and heterogeneity among respondents. The authors show how the model can be used to find market segments for bundles with heterogeneous products in multiple product categories, to estimate individual reservation prices for bundles, and to determine the optimal bundle prices for different market segments.

A General Choice Model for Bundles with Multiple-Category Products: Application to Market Segmentation and Optimal Pricing for Bundles

Although the strategy of bundling is widely used in a variety of industries (ranging from fast foods to high-technology), research in marketing has not emphasized the development of utility models for bundle choices (Russell et al. 1999). Models are lacking especially for cases in which bundles are composed of items (products) from multiple categories. Accordingly, the issues of market segmentation and target pricing for bundles composed from multiple categories also have not received much attention in the literature. This article is intended to fill that gap. The context of our work is that of a direct marketer (a firm or a reseller) offering systems of products from multiple product categories for sale to consumers. We develop and empirically test a general choice model for such systems that takes into account the interdependencies among the items in the bundle. The two distinct approaches adopted in the marketing literature to bundle choices can be classified according to the type of units for analysis: component-based and attribute-based approaches (see Figure 1).[1]

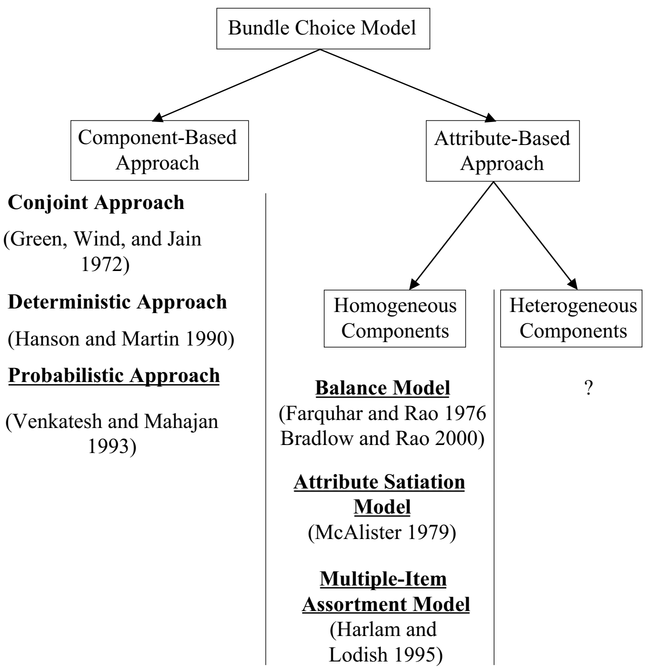
Component-based approaches refer to bundle choice models, which employ products (components) of a bundle as the ultimate unit of analysis in describing the utility of a bundle. In this stream, complementarity among the components is described directly in terms of the products included in the bundle. Three methods are employed in this approach:

*Jaihak Chung is Assistant Professor of Marketing, College of Business Administration, Sogang University, Seoul, Korea (e-mail: jjaihak@hotmail.com). Vithala R. Rao is Deane W. Malott Professor of Marketing and Quantitative Methods, S.C. Johnson Graduate School of Management, Cornell University (e-mail: vrr2@cornell.edu). The authors are grateful for the suggestions and help from Andrew Ainslie; Sachin Gupta; Greg Allenby; Peter Lenk; the three anonymous *JMR* reviewers; and attendees at the Marketing Workshop at Cornell, Marketing Science 1999 Conference, and Northeast Marketing Colloquium 2000.

[1]It is worth referring to some nonbundling studies that are helpful in studying consumers' bundle choice behaviors. There is a growing body of literature on modeling market basket choices that deal with the interdependence among multiple product categories (two examples are the articles by Manchanda, Ansari, and Gupta [1999] and Russell and Petersen [2000]). Another research stream focuses on variety-seeking behavior over time (Lattin 1987). Because of space limitations, we do not delve into this literature.

Journal of Marketing Research
Vol. XL (May 2003), 115–130

Figure 1
FIVE APPROACHES TO MODELING BUNDLE CHOICES



(1) conjoint modeling (Goldberg, Green, and Wind 1984; Green and Devita 1974; Green, Wind, and Jain 1972), (2) mixed integer linear programming (Hanson and Martin 1990), and (3) the probabilistic model (Ansari, Siddarth, and Weinberg 1996; Venkatesh and Mahajan 1993, 1997). Green, Wind, and Jain (1972) and Green and Devita (1974) apply the conjoint model to the problem of developing bundles for products such as entrees and desserts, and Wuebker and Mahajan (1999) apply the same technique for fast-food bundles. Hanson and Martin (1990) use mixed integer linear programming to determine optimal prices of a product line, and Venkatesh and Mahajan (1993) develop a probabilistic model to predict the demand for a bundle to determine optimal prices for pure bundling and mixed bundling strategies.

In contrast, the attribute-based approaches develop a model for bundle utility to specify the complementarity in terms of the attributes of the products (components). These are (1) the balance model (Farquhar and Rao 1976), (2) the attribute satiation model (McAlister 1979), and (3) the multiple-item assortment model (Harlam and Lodish 1995). Farquhar and Rao (1976) develop a utility model for bundles that incorporate interdependence among similar products by developing some measures for "balancing attributes," which we detail in the following section. McAlister (1979, 1982) initially considers the hypothesis of dependence of one item on another in an assortment and subsequently extends it for separate choice occasions with the notion of satiation due to accumulations of an attribute among the items chosen over a person's consumption history. Harlam and Lodish (1995) develop a utility model for multiple-item shopping by using the sum of variables representing within-features of items, which is, in spirit, similar to the balancing term of the balance model. In contrast to the component-based approach, the extant attribute-based models of bundle utility cannot deal with bundles of heterogeneous components that usually belong to different product categories. They are only applicable to cases in which all components in the bundle share the same attributes, though the attribute levels may differ among components.

The component-based approaches have three limitations. First, it is difficult to identify the sources of interdependence among components in a bundle when the component-based approaches are used. Second, it is impossible to use the model for forecasting the demand for bundles with new components that are not included in the model calibration; such prediction of demand is possible if bundles are defined in terms of the underlying attributes of the components. Third, in the component-based approach, the bundle space spanned by predefined components is smaller than the space generated by the attributes of these components; therefore, any optimization for bundle pricing or bundle design may lead to suboptimal solutions. Because of these factors, it is more beneficial and informative to study bundle utilities in terms of the attributes of the components, if possible at all. But the extant attribute-based approaches are limited in their applicability to studying bundles composed from heterogeneous product categories. We elaborate on this issue in the next section.

Against this background, the objective of this article is to develop a general bundle utility model based on attributes, which can be applied to any bundles consisting of both homogeneous and heterogeneous components in the pure bundling situation for bundles such as systems. This model becomes the balance model if components are comparable on all attributes (homogeneous components) and reduces to the conjoint model if components are not comparable on any attribute (heterogeneous components). Thus, it accommodates bundles with any degree of heterogeneity among components. We show how this model can be applied to finding market segments for multiple-category product bundles, an issue that has not been explored in marketing literature. Furthermore, we use this model for predicting individual reservation prices and for determining optimal bundle prices for the identified market segments.

The rest of the article is organized as follows: In the next section, we introduce a general model, called a comparability-based balance model (COBA). Then we show how to incorporate consumer preference heterogeneity and to segment customers using their preferences simultaneously. We describe the applications of this model to reservation prices and optimal bundle pricing, and then we empirically apply the model to a data set on bundle choices. We conclude with a discussion of possible directions for further research.

THE COBA

Conceptual Underpinnings

System of products. We are interested in modeling the choice of a bundle that consists of one item each from several heterogeneous product categories offered for sale by a direct marketer (a firm or a reseller). For convenience, we call this a "system." In the literature on bundling, this belongs to the pure bundling strategy. One item in a system by itself is not that useful (e.g., monitor without a computer). Therefore, the case of mixed bundling is not our focus.[2] This was not the case for Bradlow and Rao (2000), who consider assortments such as magazines (i.e., items from a homogeneous category). Although there may be cases in which consumers buy items in a system to complement their current (incomplete) systems, we focus on the purchases of complete systems. Our model is for bundles with complements under a pure bundling situation.

Attribute comparability. The balance model (Bradlow and Rao 2000; Farquhar and Rao 1976), like other attribute-based bundle choice models, assumes that the bundle components are fully comparable on all the product attributes within a bundle. Furthermore, it postulates that the attributes with nonzero influence on preference for subsets can be grouped into two classes called *nonbalancing* and *balancing* attributes. The nonbalancing attributes, which are either undesirable or desirable attributes, are the ones for which the consumer wishes to optimize (minimize or maximize) the aggregated score of the components in the bundle. The balancing attributes (equibalancing and counterbalancing) are the ones for which the consumer wants to optimize (minimize or maximize) the dispersion score of the components in the bundle. The components are equally weighted in the computation of the sums and dispersions of the attributes. The value (utility) function in the balance model is then expressed as the weighted sums and dispersions of the attributes. Relatively few simple judgments are required to determine the attribute weights in the value function. This model parsimoniously incorporates the effects of complex interactions among components into the utility of a bundle of homogeneous components.

Limitations of the balance model. For bundles of heterogeneous product categories, the balance model is limited because of the assumption that all items are comparable on all attributes. In general, for heterogeneous product bundles, some components may be unique and are noncomparable to other components. Researchers could avoid the noncomparability problem by conceptualizing higher-level attributes, which make all components comparable (Johnson 1984). However, too high a level of attribute abstraction can undermine significantly the reliability and validity of bundle evaluation data. Some concrete attributes (which are either noncomparable or only partially comparable) lose their meaning in abstraction; this may reduce the validity of responses. Additional cost of mapping concrete attributes into abstract ones can offset more than the reduction in costs of evaluating alternatives directly on concrete attributes. Furthermore, studies using abstract attributes could be of limited value for marketing decision making because of the vagueness or broadness of the resulting attribute information.

Attribute classification. To deal with this important issue of noncomparability in the balance model, we propose that consumers consider various attributes of a bundle of heterogeneous components as follows: After screening out nonessential attributes, they classify all salient attributes relevant for the bundle evaluation into one of three categories according to the degree of comparability among components: (1) fully comparable attributes, (2) partially comparable attributes, and (3) noncomparable attributes. We believe that the way consumers classify attributes is dependent on the choice situation (context); it may vary from one consumer to another, and it needs to be determined by a prior study. Our objective is not to propose a theory for such classification across all situations. This process of classification is an important avenue for further research.

We assume that people compare as many bundle components as they can to simplify the task of product evaluation if they can evaluate the products in terms of the corresponding attribute. This assumption has been implicitly employed and tested in the extant studies (Bradlow and Rao 2000; Farquhar and Rao 1976; McAlister 1979, 1982). We define *noncomparable attributes* as attributes unique to a single component of the bundle, so they are evaluated only at component levels because components cannot be compared with one another. On the contrary, *fully comparable attributes* are ubiquitous attributes that exist in any component of the bundle. A bundle can be evaluated by means of the system-level attributes because components can be compared with one another completely. Similarly, we define *partially comparable* attributes as attributes that are applicable to more than one but not all of components of the bundle. We assume that any noncomparable attribute contributes to consumer utility additively, whereas system-level attributes contribute to consumer utility additively with or without interactions according to the presence or absence of component interdependence.

As an example, consider the situation of a consumer evaluating a personal computer (PC) system consisting of a computer, a monitor, and a printer. Suppose that the consumer identified the attributes of data capacity (e.g., the capacity of hard drive), size of the monitor, speed of data processing (e.g., central processing unit [CPU] type, random-access memory [RAM] size, printer type), brand reputation, and reliability as salient for this evaluation. Brand reputation and reliability are fully comparable across components, which satisfies the comparability assumption of the balance model. The attribute of data processing speed is an example of a partially comparable attribute because computers and printers within a bundle of a PC system can be compared on this attribute but monitors cannot be.[3] There are also noncomparable attributes such as the size of a monitor and the capacity of a hard drive, which are unique to only one component of a PC system (respectively, to monitor and computer).

[2]It is equivocal what the best strategy for bundling is (Schmalensee 1984). There are conditions in which mixed bundling is optimal and others in which pure bundling is optimal. See also Venkatesh and Kamakura (2003).

[3]Technically, the speed of data processing could be viewed as a fully comparable attribute, because even monitors have some physical features regarding the speed of data processing. However, the pilot studies indicate that most respondents are not capable of evaluating monitors on this attribute, so it is unlikely to contribute to the bundle utility evaluated by them. Therefore, the speed of processing is categorized into the partially comparable attribute on which only the computer and printer can be compared.

Weighting of components. An additional consideration for heterogeneous product bundles is the necessity to weight differentially various components while computing sums and dispersions on the fully comparable and partially comparable attributes to capture interdependence among the bundle components. This is because all consumers do not treat different components as equivalent on some attributes. For example, consider the attribute of data processing speed on which both computer and printer can be compared. If the consumer has few things to print, the contribution of the printer to the bundle utility on data processing speed is little whereas that of the computer is substantial. Therefore, the components must be differentially weighted in computing sums and dispersions of attribute scores of the bundle to reflect different contributions of heterogeneous components. If some interdependence exists among components on a fully or partially comparable attribute, we capture such effects by the weighted dispersion of the attribute values across appropriate components according to the balance model (Farquhar and Rao 1976). We use a simple weighted sum for the noncomparable attributes.

The resulting model is general and applicable to evaluate and describe choices of all types of bundles (composed of both homogeneous and heterogeneous product categories). These are described in the next section.

Notation for a Bundle Utility Model

We explicitly consider the data structure collected as follows: A consumer is presented with several choice sets of multiple bundles of heterogeneous components and is asked to choose one bundle or none (no-buy option) from each choice set. Such data are collected from consumers. We allow the no-buy option in the multinomial choice set because of our interest in estimating the reservation price and the more realistic optimal price of a bundle. The corresponding experimental design ("Pick one of multiple alternatives or do not buy any of them") has been successfully employed by other studies (Louviere, Hensher, and Swait 2000). Furthermore, we consider the case of a fixed number of components in a bundle. However, the number of bundles in any choice set can vary.

We used the following symbols for developing our model:

- Q = the total number of choice sets indexed by q;
- $N(q)$ = the number of bundles in the qth choice set;
- $B(q)$ = a set of alternatives (including the no-buy option) in the qth choice set;
- $b|B(q)$ = an alternative b in the qth choice set;
- K = the number of components in a bundle, indexed by k;
- J_k = the number of products for bundling in the kth product category, indexed by j_k;
- I = the number of consumers in the sample, indexed by i;
- R = the number of covariates and intercepts;
- g = index for comparability (equals 1 for fully comparable attributes, 2 for partially comparable attributes, and 3 for noncomparable attributes);
- p_g = an attribute in the group of comparability type g, g = 1, 2, 3;
- A^g = the set of attributes characterized by comparability type g, g = 1, 2, 3;
- $Y_{ib|B(q)}$ = 1 if consumer i chooses an alternative b in the qth choice set and 0 if consumer i does not choose the alternative b in the qth choice set;
- $X^i_{j_k p_g}$ = the value of attribute p_g for product j_k given by consumer i; g = 1, 2;
- $C^b_{j_k p_3}$ = the value of the noncomparable attribute, p_3, for the item j_k in the bundle b;
- $S^{ib}_{p_g}$ = the weighted sum of the ratings of components in bundle b on attribute $p_g$with comparability type (g = 1, 2) given by consumer i;
- $D^{ib}_{p_g}$ = the weighted dispersion of the ratings of components in bundle b on attrubute p_g with comparability type (g = 1, 2) given by consumer i
- BP_b = the price of bundle b;
- RP_{ib} = the reservation price of alternative b for consumer i; and
- BV_{ib} = the value of alternative b for consumer i.

The Bundle Utility

We employ the random utility model (McFadden 1981). Considering the effects of bundle price and observed as well as unobserved attributes of bundles, the total utility of bundle b, U_{ib}, is given by

$$U_{ib} = V_{ib} + \varepsilon_{ib} \tag{1}$$

and

$$V_{ib} = BV_{ib} + \alpha_{BP}BP_b,$$

where BV_{ib} is the value of bundle b excluding the effect of price, α_{BP} (usually negative) is the coefficient for bundle price BP_b, and ε_{ib} is the error term.

We assume that the value of a bundle, BV_{ib}, consists of products (components) from K different product categories. Suppressing the subscript i, we can write BV_b as (part of the deterministic component of the utility of bundle b by consumer i)

$$BV_b = \alpha_0 + \sum_{p_1 \in A^1} \left[\beta_{p_1} S^b_{p_1} + \gamma_{p_1} D^b_{p_1}\right] + \sum_{p_2 \in A^2} \left[\beta_{p_2} S^b_{p_2} + \gamma_{p_2} D^b_{p_2}\right] + \sum_{p_3 \in A^3} \alpha_{p_3} C^b_{p_3}, \tag{2}$$

where α_0 is the intercept parameter, and β_{p_g} and γ_{p_g} are coefficients for nonbalancing covariates, $S^b_{p_g}$, and balancing covariates, $D^b_{p_g}$, with comparability type g (equals 1 and 2, respectively, for fully comparable and partially comparable attributes). The parameter α_{p_3} is the coefficient for the noncomparable attribute p_3 for bundle b with the level $C^b_{p_3}$, which can be a physical value or consumer i's perceived value. According to the researcher's purpose, dummy variables can be used to represent $C^b_{p_3}$ instead of the scaled values. Then, its coefficient can be interpreted as the partworths of the corresponding attributes in the conjoint model.

Note that a positive (or negative) β_{p_g} implies that the attribute p_g is desirable (undesirable), and a positive (or negative) value of γ_{p_g} implies that the attribute p_g is counterbalancing (or equibalancing). The terms $S^b_{p_g}$ and $D^b_{p_g}$ are derived from respondents' evaluations, $X_{j_k p_g}$, of product j_k on the attribute p_g as shown:

(3) $$S^{b}_{p_g} = \sum_{k=1}^{K} w_k \sum_{j_j=1}^{J_k} 1(j_k, p_g, b) X_{j_k p_g}, \text{ where } g = 1, 2;$$

(4) $$D^{b}_{p_g} = \sum_{k=1}^{K} w_k \sum_{j_j=1}^{J_k} 1(j_k, p_g, b)\left(X_{j_k p_g} - \overline{X}^{b}_{p_g}\right)^2,$$

where $g = 1, 2;$ and

(5) $$\overline{X}^{b}_{p_g} = \frac{\sum_{k=1}^{K} \sum_{j_k=1}^{J_k} 1(j_k, p_g, b) X_{j_k p_g}}{\sum_{k=1}^{K} \sum_{j_k=1}^{J_k} 1(j_k, p_g, b)}, \text{ where } g = 1, 2,$$

$$\text{where } 1(j_k, p_g, b) = \begin{cases} 1 & \text{if the jth item in category k is included in bundle b and can be evaluated on attribute } p_g, \\ 0 & \text{otherwise,} \end{cases}$$

and w_k is an importance weight of any component in the kth product category for bundling.

Basically, $S^{b}_{p_g}$ for bundle b is a weighted sum of the ratings of relevant components for bundle b on attribute p weighted by the corresponding category importance w_k, and $D^{b}_{p_g}$ for bundle b is a weighted dispersion of the ratings of relevant components in bundle b and the mean rating of bundle b on attribute p_g, weighted by w_k.[4] We obtain w_k by questioning respondents on the importance of each component to the system. The calculations of $S^{b}_{p_g}$ and $D^{b}_{p_g}$ for fully comparable (g = 1) and partially comparable (g = 2) attributes are different. For the fully comparable attribute (g = 1), ratings of components from all categories are used in the calculation of $S^{b}_{p_g}$ and $D^{b}_{p_g}$. For the partially comparable attribute (g = 2), only the ratings of components that can be compared are used. For example, suppose that only the computer and printer (not the monitor) can be compared by the consumers in the sample on the attribute of data processing speed; this attribute belongs to the second type (partially comparable). For this attribute, the ratings of a subset of components (e.g., computer and printer) are used in the calculation of $S^{b}_{p_g}$ and $D^{b}_{p_g}$ for bundle b. In contrast, the $S^{b}_{p_g}$ and $D^{b}_{p_g}$ terms for the fully comparable attribute, such as brand reputation for bundle b, are calculated from the ratings of all components forming the bundle b (e.g., a monitor, a computer, and a printer in a bundle b).

The model uses a more general form of covariates involving nonbalancing and balancing attributes by allowing for the different contributions of each component on the same attribute to the bundle utility: $\Sigma^{K}_{k=1} w_k = 1$.[5] The category importance weights play an important role in describing the unique preference structure of system bundle compared with an assortment consisting of magazines or candy bars. For example, shoppers for PC systems might value computers more than other components such as printers and monitors. For system bundles such as computer systems and stereo systems, each component inherently contributes differently to the bundle utility because of its functional heterogeneity. The model is so flexible that it can also handle assortment choices without any modification. The standard balance model is also a special case of this model.

Incorporating Heterogeneity and Market Segmentation

To find market segments by means of consumers' choice behaviors toward bundles and to incorporate consumer heterogeneity, we employ a mixture distribution modeling approach (Allenby, Arora, and Ginter 1998; Diebolt and Robert 1994) in a hierarchical Bayesian framework. Each consumer is assumed to belong to a particular segment exclusively among a finite number, F, of heterogeneous segments to describe heterogeneity across segments (Kamakura and Russell 1989). In this finite mixture model, not only are segments allowed to differ from one another, but also consumers within a segment are allowed to differ in their preferences toward bundles.

We employ the finite mixture distribution modeling for several reasons. First, we employ latent-class modeling to find market segments based on consumers' idiosyncratic preferences on product attributes. Second, we allow for heterogeneity within classes because there is some evidence that within-segment heterogeneity can be too large to ignore in model estimation (Allenby, Arora, and Ginter 1998). The mixture distribution model is general and includes the latent-class model and standard random-coefficient choice model as special cases; we therefore are able to determine the degree of heterogeneity across and within segments and assess the need for this general model. Our empirical study as described in the estimation section shows the necessity of allowing for the heterogeneity within classes in a latent-class structure. Third, the finite mixture distribution is flexible enough to estimate any nonstandard shape of preference heterogeneity distribution such as nonunimodal or skewed. In this sense, the finite mixture distribution modeling can be interpreted as a parsimonious alternative to nonparametric modeling. Therefore, the individual-level estimates obtained from the finite mixture distribution modeling can be employed to obtain more-accurate estimates of optimal prices and individual-specific reservation prices.

In light of this discussion, we can rewrite the deterministic part of the utility of bundle b for consumer i, given that he or she belongs to segment f, $V_{ib|f}$, as

(6) $$V_{ib|f} = \alpha^{fi}_{0} + \sum_{p \in A^1} \left[\beta^{fi}_{p_1} S^{ib}_{p_1} + \gamma^{fi}_{p_1} D^{ib}_{p_1}\right] + \sum_{p \in A^2} \left[\beta^{fi}_{p_2} S^{ib}_{p_2} + \gamma^{fi}_{p_2} D^{ib}_{p_2}\right] + \sum_{p_3 \in A^3} \alpha^{fi}_{p_3} C^{b}_{p_3} + \alpha^{fi}_{BP} BP_b.$$

Note that all preference parameters in this equation are individual-specific within a segment. Therefore, V_{ib} is

(7) $$V_{ib} = \sum_{f=1}^{F} \psi_{fi} V_{ib|f},$$

[4]If an attribute does not change, the dispersion will naturally be zero; therefore, there is an implied restriction in the model that D is not zero when gamma is in the model.

[5]We can use a more general importance term such as categorywise or attributewise importance, w_{kp}. In our study, we use categorywise importance, w_k, which is helpful in reducing the number of questions asked of respondents.

where ψ_{fi} is the probability of consumer i belonging to segment f.

Estimation Method

To allow for the correlation among the error terms of bundle utilities due to the unobserved similarity among bundles against the no-buy option, we employ the nested logit model. The cumulative density function of error terms, $\{\varepsilon_{ib}\}$, is given as follows (McFadden 1981):

$$F\left(\{\varepsilon_{ib}\}_{b=0}^{N(q)}\right) = \exp\left\{-\exp\{-\varepsilon_{i0}\} - \left[\sum_{b=1}^{N(q)} \exp\left(-\frac{\varepsilon_{ib}}{\rho_1}\right)\right]^{\rho_{i1}}\right\}$$

$$0 < \rho_{i1} \le 1.$$

All error terms of bundles in the qth choice set, $\{\varepsilon_{ib}\}_{b \neq 0}$, are correlated, whereas the error term of the utility of the no-buy option, ε_{io}, is independent of the other error terms. Assuming that the consumer chooses an alternative with the maximum random utility among all the alternatives in a choice set, the likelihood can be given as

$$L = \prod_i \prod_{q=1}^{Q} \sum_{f=1}^{F} \left\{ P_{i0|B(q),i \in f}{}^{D_{i|B(q)}} \prod_{b=1}^{N(q)} \left[P_{ib|b \neq 0,B(q),i \in f}\left(1 - P_{i0|B(q),i \in f}\right)\right]^{D_{bi|B(q)}} \psi_f \right\},$$

$$\text{where } D_{ib|B(q)} = \begin{cases} 1 & \text{if consumer i chooses alternative b in the qth choice set,} \\ 0 & \text{if consumer i chooses any bundle in the qth choice set,} \end{cases}$$

and $b = 0$ indicates the no-buy option.

Given that consumer i does not choose the no-buy option in a choice set q and belongs to segment f, the probability of his or her choosing bundle b can be derived as

$$P_{ib|b \neq 0,B(q),i \in f} = \frac{V_{ib|i \in f}}{\sum_{b' \neq 0,\ b' \in B(q)} \exp\left(V_{ib'|i \in f}\right)} \text{ for } b \in B(q).$$

Furthermore, the probability of consumer i's choosing the no-buy option (b = 0) in a choice set q can be derived as (McFadden 1981)

$$P_{i0|B(q),i \in f} = \frac{1}{1 + \exp\left(\rho_{i0} + \rho_{i1} IV_{B(q)}\right)},$$

where $\rho_{i0} = -U_{i0}$, ρ_{i1} is an individual-specific correlation coefficient, and

$$IV_{B(q)} = \ln\left[\sum_{b=1}^{N(q)} \exp\left(\frac{V_{ib}}{\rho_{i1}}\right)\right].$$

Note that ρ_{i1} and ρ_{i0} are individual-specific, whereas the vector $\phi_i = (\alpha_i, \beta_i, \gamma_i)'$ is specific to a consumer within a segment. We use hierarchical Bayesian methods in estimating the parameters of this model (Poirier 1996). We give technical details in the Appendix.

APPLICATION: ESTIMATION OF RESERVATION PRICES AND OPTIMAL PRICES

In general, a consumer's reservation price for bundle b depends on the set of alternatives in the choice set, B(q), because it is the bundle price at which he or she is indifferent between buying bundle b and choosing the other alternatives in a choice set.[6] According to the definition of reservation price, the utility of bundle b in a choice set B(q), when the price for bundle b is equal to its reservation price, is equal to the utility of the best alternative among the remaining alternatives in the choice set, $U_{max|B(q)-b}$, where $B(q) - b$ denotes the set of bundles excluding the target bundle b in the qth choice. Denoting the reservation price for the target bundle b given the qth choice set as $RP_{b|B(q)}$, we get the expression

$$U_b\left(RP_{b|B(q)}\right) = U_{max|B(q)-b}, \tag{8}$$

$$\text{where } U_{max|B(q)-b} = \text{Max}\left[\underset{\substack{b' \in B(q)\\ b' \neq b\\ b' \neq 0}}{\text{Max}}\{U_{b'}\}, \frac{U_{io}}{\rho_1} - \sum_{\substack{b' \in B(q)\\ b' \neq b\\ b' \neq 0}} \exp\left(\frac{U_{b'}}{\rho_1}\right)\right].$$

Therefore, consumer i's reservation price for bundle b in the qth choice set is given by

$$RP_{b|B(q)} = -\frac{BV_b}{\alpha_{BP}} - \frac{\varepsilon_b}{\alpha_{BP}} + \frac{U_{max|B(q)-b}}{\alpha_{BP}}.$$

Because the model parameters and error terms are random variables, the estimated reservation price has a probability distribution. Therefore, we can estimate the expected reservation price by using the following equation:

$$E\left(RP_{b|B(q)}\right) = -E\left[\frac{\alpha_0}{\alpha_{BP}} + \sum_{p_1 \in A^1}\left(\frac{\beta_{p_1}}{\alpha_{BP}} S_{p_1} + \frac{\gamma_{p_1}}{\alpha_{BP}} D_{p_1}\right) + \sum_{p_2 \in A^2}\left(\frac{\beta_{p_2}}{\alpha_{BP}} S_{p_2} + \frac{\gamma_{p_2}}{\alpha_{BP}} D_{p_2}\right) + \sum_{p_3 \in A^3} \frac{\alpha_{p_3}}{\alpha_{BP}} C_{p_3}\right] + \frac{E\left(U_{max|B(q)-b}\right)}{\alpha_{BP}}. \tag{9}$$

We first draw a sample of parameter vectors from their posterior distribution through Markov chain Monte Carlo (MCMC) simulation and compute the right-hand side of Equation 9 for each draw. We use the average of these calculated values as the estimate of the reservation price for a consumer for a specific bundle.

The implications of Equation 9 are evident if it is rewritten as $E(RP_b) = BP_c + (BV_b - BV_c)/(-\alpha_{bp})$, where c refers

[6]Kohli and Mahajan (1991) estimate reservation price in a similar manner, but they ignore the no-buy option in their framework, which could result in the overestimation of reservation price.

to the best of the remaining bundles in the choice set and b refers to the target bundle. Here, the best bundle's price is BP_c. It is clear how the reservation price for a target bundle depends on the valuation of the target bundle compared with the best of the remaining bundles in the choice set. Because the sign of the price coefficient (α_{bp}) is negative, a consumer's reservation price will be higher (lower) than the price of the best bundle when the target bundle is valued higher (lower) than the best bundle. The reservation price for the target bundle will be the same as the best bundle's price when their evaluations are equal.

Furthermore, if the set of other bundles in the choice set is held constant, Equation 9 indicates that the reservation price for the target bundle will be higher for attributes regarded as desirable (i.e., $\beta_{pg} > 0$ or $\alpha_{p_3} > 0$) and lower for attributes regarded as undesirable (i.e., $\beta_{pg} < 0$ or $\alpha_{p_3} < 0$). Similarly, the reservation price of the bundle will be higher for attributes regarded as counterbalancing or complementary (i.e., $\gamma_{pg} > 0$) and lower for attributes regarded as equibalancing or similar (i.e., $\gamma_{pg} < 0$). Another consequence of this formulation is that the reservation price of a bundle is simply the sum of the reservation prices of the items in the target bundle when there is no counterbalancing or equibalancing in the bundle evaluation (i.e., $\gamma_{pg} = 0$ for all p). Thus, the model explicitly considers the interdependencies among items as expressed in terms of their attributes.

We can determine the optimal price for a particular bundle b from a choice set, which maximizes the profit contribution from the bundle. Let us assume that the firm's cost for producing the bundle is constant, denoted by BC_b. The expected profit from the offer of the bundle b using the pure bundling strategy is

$$\text{Expected profit} = M\,(BP_b - BC_b)\,\Sigma^I_{i=1}P_{ib},$$

where M is a factor that projects the sample of I consumers to the potential market as a whole, and P_{ib} is the probability of consumer i choosing the bundle b. In this formulation, the expected profit and P_{ib} depend on the set of bundles that appears along with bundle b.

The optimal price BP_b* for bundle b can be determined by maximizing expected profit with respect to BP_b. The first-order conditions yield the following equation for determining the optimal bundle price, BP_b*:

$$(10)\qquad BP^*_b = BC_b - \frac{\sum_{i=1}^{I} P^*_{ib}}{\sum_{i=1}^{I} \left.\frac{\partial P_{ib}}{\partial BP_b}\right|_{BP_b = BP^*_b}},$$

where (*) denotes the values of P_{ib} at the optimal price BP_b*, and mean values are used for preference parameters in P_{ib}. The second-order condition for the global maximum for optimal prices is satisfied when the bundle price for b, BP_b, is greater than or equal to the cost of bundle b, BC_b, and the price sensitivity α^i_{BP} is negative for all i. Because there is no closed-form solution for this optimization problem, we use numerical optimization to determine optimal prices. It is worth noting that the optimal price obtained from this model under pure bundling can be suboptimal because pure bundling may be a suboptimal strategy when components for bundles are not strongly complementary. For details on the conditions in which pure bundling strategy can be suboptimal, see Venkatesh and Kamakura (2003).

EXPERIMENTAL STUDY

Study Design

Context. The context of our study is that of undergraduate students' purchase decisions for PC systems consisting of a monitor, a computer, and a printer as a bundle. Students were asked to imagine that they were planning to purchase a PC system offered for sale by a local retailer and that they had enough funds available for purchase. They were asked to choose the no-buy option or to buy one of the bundles given in 18 independent choice tasks.

A sample of 136 undergraduate students participated in this experiment. We used PC systems for our study because they are a good example of bundles with multiple components in which components are heterogeneous in terms of attribute structure, such as functionality and physical features, and are relevant to the respondent sample. Furthermore, bundling a computer, a monitor, and a printer is a predominant marketing practice in the PC industry (Wilson, Weiss, and John 1990).

Identification of attributes/physical features. We conducted two pilot studies for identifying the critical attributes regarding PC systems, by using interviews and questionnaires with sample sizes of 18 and 27, respectively, in May 1997 and February 1998. We conducted pilot studies by directly questioning respondents about what physical features and attributes they consider for PC system bundle evaluation and choice. On the basis of these pilot studies, we selected four physical features (i.e., monitor size, CPU type, RAM, and hard drive capacity)[7] and three attributes (i.e., brand reputation, data processing speed, and reliability) as important for customers' choice of PC systems. The CPU type and RAM features are naturally mapped into the attribute of data processing speed, leaving only two physical features to be included in the model. In the pilot studies, we sought information on the degree of comparability of each physical feature and attribute. For each feature and attribute, respondents rated all the components of the three product categories (computers, printers, and monitors) on whether the components are fully comparable, partially comparable, or not comparable. Using this information, we ascertained that all consumers adopt the same classification of attributes for a bundle in the case of PC systems.

Stimuli. We used three brands for Monitors (Sony 21″, ViewSonic 17″, and ViewSonic 15″), three brands for PCs (Dell Dimension, HP Vectra, and eMachines eTower), and two brands for Printers (HP Laser printer 1100 and Epson Inkjet printer Stylus 300) in the development of bundles for our study. We used the sum of prices of each bundle component as the price of the corresponding bundle. Specific descriptions and prices of the various components used are

[7]The pilot studies were consistent in the finding that the resolution of monitors and the speed of the CD-ROM are not critical features for the evaluation of PC systems compared with other features. For example, all 45 respondents participating in pilot studies were not able to evaluate the resolution of monitors (e.g., the number of pixels for a monitor's resolution). Furthermore, we found that they were insensitive to the levels of CD-ROM speed provided in the study. Accordingly, we did not present information on the resolutions of monitors in the main study, and we used the same level of CD-ROM speed for all computers to reduce the complexity of bundle profiles.

shown in Table 1. The prices for all components used in our study had been obtained from a major Internet search engine specialized on computer hardware and a major e-tailer.[8] For this task, all 18 possible combinations of bundle components (three monitors, three computers, and two printers), showing real street prices (retail prices) for each component, are generated as bundles.

Choice sets. We designed 18 choice sets, each containing three to five alternatives bundles including the no-buy option. Five choice sets contain two bundles, ten choice sets contain three bundles, and three choice sets contain four bundles. We randomized bundles in each choice task.

Data collected. For each choice set, respondents indicated which bundle or no-buy option they would choose. They indicated whether the three components (monitor, computer, and printer) are comparable on each of three attributes (brand reputation, reliability, and data processing speed). They also rated each component on its importance in the bundle using a ten-point (1–10) scale. They were asked to rate the eight components (three each for monitors and computers and two for printers) on several attributes on the basis of comparability types on a 1–10 scale. In addition, respondents indicated their most preferred bundle and how much they would be willing to pay for it. We used four different versions of questionnaires in which the questions were presented in different orders. Data were collected by a self-administered questionnaire.

Attribute/physical feature classification. Using the information collected in both the pilot studies and the main study, we arrived at a classification of the three attributes (brand reputation, reliability, and data processing speed) and the two physical features (monitor size and hard drive capacity) as shown in Table 2. We used this classification in the estimation of our model.

Model Estimation

Number of parameters. We estimate the model developed previously using information on the fully comparable attributes (brand reputation and reliability), the partially comparable attribute (speed), and the two noncomparable attributes (monitor size and the capacity of hard drive). There are nine parameters (eight for attributes and physical features and one for price) in the model that are estimated for different numbers of segments. Furthermore, we retain heterogeneity within each segment.

Estimation and validation. We used data from 14 randomly selected purchase decisions (choice sets) per respondent for estimation (1904 observations) and data from the remaining four purchase decisions per respondent for validation (544 observations).

[8]See http://www.zdnet.com/computershopper and http://www.pcmall.com.

Table 1

EXPERIMENTAL STUDY: PC SYSTEM

Bundle Component		Physical Features				
		Monitor Size	*Price*			
Monitor	Sony	21 inch	$400			
	ViewSonic	17 inch	$180			
	ViewSonic	15 inch	$100			
		CPU (Pentium)	*Ram*	*Hard Disk*	*CD-ROM*	*Price*
PC	Dell Dimension	1000 MHz	256 MB	40 GB	40X	$2,200
	HP Vectra	553 MHz	128 MB	20 GB	40X	$1,400
	EMachines eTower	333 MHz	64 MB	20 GB	40X	$ 510
		Price				
Printer	HP Laser Printer 1100	$310				
	Epson Inkjet Printer Stylus 300	$130				

Table 2

THE CLASSIFICATION OF PHYSICAL FEATURES

Component	*Attributes*	*Noncomparable*	*Partially Comparable*	*Fully Comparable*
Monitor[a]	Brand Reliability Monitor size	Monitor size		Brand reputation Reliability
PC[b]	Brand Reliability CPU Hard drive	Data capacity	Data processing speed	Brand reputation Reliability
Printer	Brand Reliability Printing speed		Data processing speed	Brand reputation Reliability

[a]The information on the resolution of the cathode ray tube is not provided to the respondents.
[b]The speeds of CD-ROMs are given at the same level, 40X.

Table 3

IN-SAMPLE AND OUT-OF-SAMPLE FITS FOR THE COBA MODEL: LOG MARGINAL LIKELIHOOD AND HIT RATE

Number of Classes	*In-Sample Fit* *Log Marginal Likelihood*[a]	*Out-of-Sample Fit (Hit Rate) for the Choice-Set Size of* *3*	*4*	*5*
1	–820.55	.67	.59	.51
2	–701.14	.74	.77	.61
3	–769.66	.70	.64	.57

[a]See footnote 9.

We used the log marginal likelihoods of models with the different numbers of latent segments as the criterion for estimation; the log marginal likelihood is computed as the logarithm of the harmonic means of the likelihood values using estimated parameter samples drawn by Gibbs sampling (Newton and Raftery 1994). We used hit rates to validate the model for the holdout sample.

Segment selection. Using log marginal likelihoods as an in-sample fit criterion and hit rates as a predictive validity criterion, we determined the model with two segments to be the best model. Because the marginal likelihood is penalized by model complexity (Chib and Jeliazkov 2000; Meliă and Heckerman 1998), it increases and then decreases at some point as the number of segments increases.[9] The two-segment model gave us meaningful managerial interpretation and outperforms other models, as is shown in Table 3. We could not estimate more than a four-segment model because there were some segments at some iteration periods to which no subjects were assigned.

Parameter estimates. The means and standard deviations of the posterior distributions for the preference parameters $\phi_i = (\alpha_i, \beta_i, \gamma_i)'$ of our model, which we label as Model 7 in the following model comparison section, are shown in Table 4. The mean of the dissimilarity coefficient (ρ_{i1}) is .91 with a standard deviation of .119. The mean of the constant (ρ_{i0}) is .0705 with a standard deviation of .0251. The last column in Table 4 shows the expected signs of preference parameters. We expect the sign for price to be negative and the sign of all nonbalancing attributes (desirable) to be positive. We also expect that the signs of the dispersion coefficients for brand reputation, reliability, and data processing speed will be negative because of their equibalancing effects on the utility of a system bundle. The signs of the average parameter values follow our expectations for both segments.

We find that the nonbalancing and balancing effects for the fully comparable attributes (i.e., brand reputation and reliability) are in the right direction and are significant for both segments. We also find that the nonbalancing and balancing effects of the partially comparable attribute (speed) and the effect of the noncomparable attributes (monitor size and capacity of hard drive) are significant for the two segments. Note that the balance model cannot incorporate the important information as to which attribute is partially comparable or noncomparable. The magnitude of the attribute effects varies according to customer segment and type of attribute. These results indicate that respondents consider the notion of balance in evaluating PC systems. However, the sensitivities for the balancing attributes differ by segment and type of attribute.

Figure 2, Panel A, shows density estimation plots of Segments 1 and 2 to describe the heterogeneity of the price sensitivity parameter in our model. We find that the two segments differ in price sensitivity. The density estimation plots of all respondents for preference heterogeneity for five parameters of our model are shown in Figure 2, Panels B through E. We used the individual means of 3000 draws generated from posteriors of individual preference parameters in these figures. Note that most of them do not show the form of symmetric unimodal distribution, which reinforces the necessity of using mixture distribution in our analysis.

Table 5 shows the degree of consumer heterogeneity within and across segments. We use the average of the within and across sum of squares of the individual posterior parameter values for each iteration as the measure of heterogeneity. Heterogeneities across segments for some attributes such as capacity of hard drive and price are notably high, whereas those for the other attributes such as data processing speed and brand reputation are low. Again, they show multimodality of consumer heterogeneity distributions for some attributes such as monitor size and brand reputation, in support of the usefulness of the finite mixture modeling approach. This heterogeneity decomposition in Table 5 also shows which attributes are more important in segmenting markets, so this method can be used in an exploratory analysis to determine the attributes for which a mixed-effect latent-class model is appropriate. This decomposition method is a useful tool to analyze and measure the degree of heterogeneity, which has not hitherto been introduced in the marketing literature.

Segment interpretation. In our study, we collected information on characteristics of respondents such as sex, age, monthly expenditure, computing time, and computing proficiency. On the basis of these data and the parameter estimates in Table 4, we can interpret the two segments as follows: Segment 1, accounting for 39% of the sample, can be

[9]Suppose that we consider a set of K models $\{\Theta_k\}$ that reflects competing hypotheses about the data set y. Then, the marginal likelihood of a model Θ_k, $m(y|\Theta_k)$, is given by $m(y|\Theta_k) = \int f(y|\Theta_k, \phi_k)\pi(\phi_k|\Theta_k)d\phi_k$, where $f(y|\Theta_k, \phi_k)$ is a density function of y conditional on model k and the corresponding model parameter ϕ_k, and $\pi(\phi_k|\Theta_k)$ is the prior for the corresponding model parameter ϕ_k. Note that the marginal likelihood consists of not only likelihood but also priors. Because the marginal likelihood is the normalizing constant of the posterior density, the marginal likelihood can be expressed as

$$m(y|\Theta_k) = f(y|\Theta_k, \phi_k)\pi(\phi_k|\Theta_k)/\pi(\phi_k|y, \Theta_k).$$

Evaluating the right-hand side of this equation at some appropriate point ϕ_k^* and taking logarithms, log marginal likelihood can be given by

$$\ln\left[m\left(y|\Theta_k\right)\right] = \ln\left[f\left(y|\Theta_k, \phi_k^*\right)\right] + \ln\left[\pi\left(\phi_k^*|\Theta_k\right)\right] - \ln\left[\pi\left(\phi_k^*|y, \Theta_k\right)\right].$$

This equation shows that the marginal likelihood is determined not only by the simple likelihood but also by two other factors, the prior $\pi(\phi_k^*|\Theta_k)$ and the posterior $\pi(\phi_k^*|y, \Theta_k)$ evaluated at a parameter point ϕ_k^*. The value obtained from evaluating the posterior, $\pi(\phi_k^*|y, \Theta_k)$, at a point ϕ_k^*, can increase as the number of model parameters increases. In other words, the last term can play a role of penalizing the marginal likelihood as more model parameters are allowed. Therefore, marginal likelihood can increase and decrease as the number of model parameters increases. Marginal likelihood is somewhat similar to penalized likelihoods such as Bayesian information criterion. For example, Bayesian information criterion is an approximation of marginal likelihood (Chickering and Heckerman 1997).

Table 4

MEANS AND STANDARD DEVIATIONS OF PARAMETERS FOR THE COBA MODEL, $\phi_i = (\alpha_i, \beta_i, \gamma_i)'$

Attribute Type	*Attribute*	*Parameter*	*Segment 1*[a]	*Segment 2*[a]	*Expected Sign of Parameter*
	Price	Price	–.0109** (.0045)	–.0212** (.0091)	–
Fully comparable	Brand reputation	S1	1.2309 (.4509)	1.3565 (.2955)	+
		D1	–.5061* (.5110)	–.6979 (.9456)	–
	Reliability	S2	1.1028** (.2979)	.8379* (.3584)	+
		D2	–.4276** (.2021)	–.2207* (.1615)	–
Partially comparable	Speed	S3	2.4269 (.7598)	2.5251* (.7570)	+
		D3	–1.1432* (.5417)	–1.1599* (.7798)	–
Noncomparable	Monitor size	C1	.5559 (.2599)	.8320* (.5299)	+
	Capacity of hard drive	C2	.2144 (.2010)	.7339* (.4319)	+
		Membership probability	.3897* (.1355)	.6103* (.1355)	
		Intercept	4.2912 (1.7388)	6.4566 (4.3099)	

*$p < .05$.
**$p < .01$.
[a]The numbers in parentheses are standard deviations of the corresponding parameters obtained by averaging the square roots of diagonal elements of the covariance matrix T_f draws.
Notes: Estimated probabilities of the corresponding parameter being inconsistent with the expected signs are used as an analog of *p*-values.

characterized as consumers who are more familiar with computers,[10] less sensitive to price, and more concerned about reliability than are consumers in Segment 2 in evaluating the values of PC systems. In contrast, Segment 2, with a larger proportion of the sample (61%), can be characterized as less familiar with computers, more sensitive to price, and more concerned about attributes such as capacity of hard drive and monitor size than are consumers in Segment 1. The members of Segment 2 are much less concerned about experience attributes such as reliability of bundle components. Although other differences exist, we may succinctly describe Segment 1 as performance oriented and Segment 2 as economy oriented.

Model Comparison

Table 6 lists six models (Models 1–6) according to types of information used, the number of segments, and the existence of preference heterogeneity. For completeness, we show our COBA with two latent classes estimated previously as Model 7.

Model 1: This model treats the components of the bundle as noncomparable on all attributes. This is akin to a conjoint utility model. Accordingly, it restricts all parameters of fully and partially comparable attributes to zero (all $\beta_{1p} = \beta_{2p} = 0$, all $\gamma_{1p} = \gamma_{2p} = 0$) and ignores heterogeneity. Instead, it uses some additional information (physical features: CPU speed, RAM size, and a dummy for the types of printer) for the fair comparison of the performance of conjoint model with the COBA models. The preference parameters in Model 1 are individual specific.

Model 2: This is the original balance model that does not incorporate balance effects. It uses only the nonbalancing effects of fully comparable attributes without heterogeneity and the latent-class structure. Accordingly, it restricts parameters for the balancing effects of fully comparable attributes to zero (all $\gamma_{1p} = 0$) and parameters for partially and noncomparable attributes to zero (all $\beta_{2p} = 0$, all $\gamma_{2p} = 0$, and all $\alpha_{kp} = 0$).

Model 3: This is the original balance model. It does not include any partially comparable or noncomparable attributes and ignores heterogeneity. Accordingly, it restricts parameters for partially and noncomparable attributes to zero (all $\beta_{2p} = 0$, all $\gamma_{2p} = 0$, and all $\alpha_{kp} = 0$).

Model 4: This is the COBA model that ignores any noncomparable attributes and any heterogeneity among the sample respondents. It restricts all $\alpha_{kp} = 0$.

Model 5: This is the COBA model that assumes no heterogeneity.

Model 6: This is the COBA model for the sample as a whole (or one latent class). It incorporates heterogeneity among the sample respondents.

[10]The means of computing time and computer proficiency for Segment 1 are 23 (hours per week) and 7.12 (on a ten-point scale), and those for Segment 2 are 15 (hours per week) and 5.88 (on a ten-point scale).

Model 7: This is the same as Model 6 but for two latent classes of respondents.

We note that Bradlow and Rao's (2000) model for assortments is also a special case of Model 6, which allows for heterogeneity in model parameters but does not include any partially comparable or noncomparable attributes. We estimated all models using the nested multinomial logit framework.[11] We compare Models 1 to 7 in terms of in-sample and out-of-sample fits to determine how helpful the inclusion of the two factors (partially comparable and noncomparable attributes and heterogeneity within latent classes) are in predicting bundle choices (see Table 7 for details). We used two criteria: log marginal likelihood for in-sample fits of models and hit rates for out-of-sample fits (proportion of correctly predicted choices) by the size of the choice set (3, 4, or 5).

We estimate Model 1 (conjoint model) using all available physical features as attributes.[12] This is somewhat different from simply ignoring the fully comparable and partially comparable attributes.

[11]The nested logit model shows significant improvement by 11% (average) in terms of predictive validity (hit rates) over the nonnested logit model. We acknowledge the suggestions from reviewers on this revision.

[12]Additional physical features are CPU clock speeds, RAM sizes, and a dummy variable for Inkjet or laser printer. For details, see Table 1.

Figure 2

ESTIMATED HETEROGENEITY (POSTERIOR DISTRIBUTIONS) FOR PREFERENCE PARAMETERS

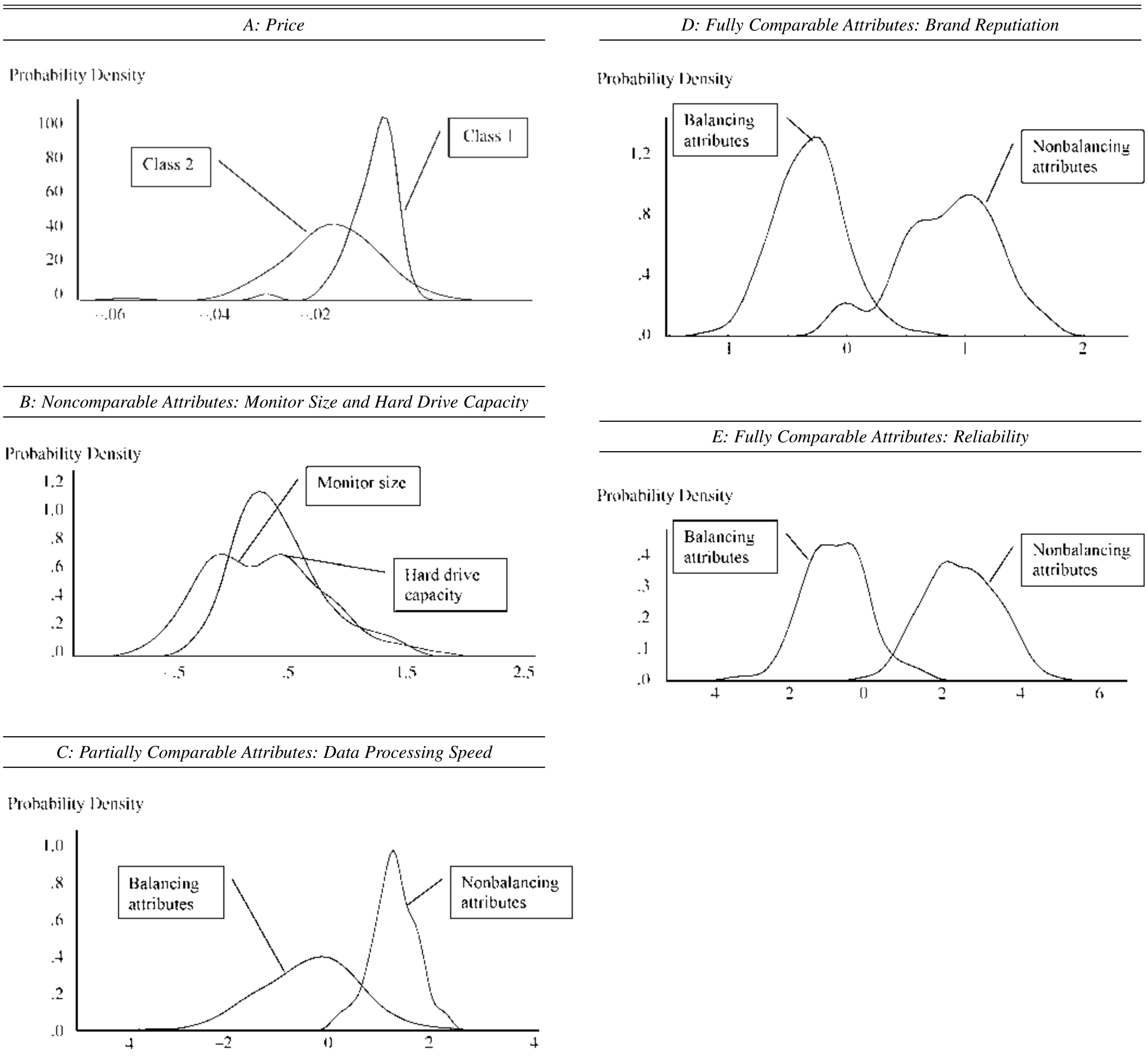


Table 5

HETEROGENEITY DECOMPOSITION FOR THE COBA MODEL

Attribute Type	*Attribute*	*Parameter*	*Heterogeneity Within Segment 1*	*Heterogeneity Within Segment 2*	*Heterogeneity Across Segments 1 and 2*	*Heterogeneity*[a]
	Price	Price	.0011 (10.51)[b]	.0068 (67.96)	.0022 (21.54)	.0100 (100%)
Fully comparable	Brand reputation	S1	10.57 (57.95)	7.16 (39.26)	.51 (2.80)	18.24 (100%)
		D1	13.58 (15.41)	73.32 (83.23)	1.19 (1.35)	88.08 (100%)
	Reliability	S2	4.61 (26.49)	10.53 (60.49)	2.27 (13.02)	17.42 (100%)
		D2	2.12 (37.60)	2.14 (37.88)	1.38 (24.52)	5.65 (100%)
Partially comparable	Speed	S3	30.02 (38.82)	46.98 (60.77)	.31 (.40)	77.31 (100%)
		D3	15.26 (23.43)	49.86 (76.56)	.01 (.01)	65.13 (100%)
Noncomparable	Monitor size	C1	3.51 (12.11)	50.35 (79.39)	.045 (8.50)	29.01 (100%)
	Capacity of hard drive	C2	2.10 (8.04)	15.30 (58.55)	8.73 (33.41)	26.13 (100%)
		Intercept	28.38 (1.67)	1523.15 (89.43)	151.67 (8.90)	1703.20 (100%)

[a]These entries are sum of squared error within each of the two segments, between the two segments, and total.
[b]These are rowwise percentages of the sum of squared error values.

Table 6

ALTERNATIVE MODELS COMPARED

Model	*Latent Class*	*Unobserved Heterogeneity (Random Effects)*	*Interdependence Among Components*	*Fully Comparable Attribute*	*Partially Comparable Attribute*	*Noncomparable Attribute*
Model 1: Conjoint model (Green, Wind, and Jain 1972)	1	√				√
Model 2: Balance model without balancing attributes	1			√		
Model 3: Balance model (Farquhar and Rao 1976)	1		√	√		
Model 4: COBA without noncomparable attributes	1		√	√	√	
Model 5: COBA without heterogeneity	1		√	√	√	√
Model 6: COBA with heterogeneity	1	√	√	√	√	√
Model 7: COBA within two segments	2	√	√	√	√	√

The results show that Model 7 outperforms the other models, including the conjoint model and the balancing model in both in-sample and out-of-sample fits. Not surprisingly, the simplest model, Model 2, is the worst among these models. Compared with Model 3 (which does not include any partially comparable or noncomparable attributes), Model 5 (the homogeneous COBA model) enhances in-sample fit by 19% (–1012.33 versus –1199.53) and hit rates by 21%, 16%, and 19% for the three choice-set sizes (3, 4, and 5), respectively; this suggests that the partially comparable attribute (speed) and the noncomparable attributes (monitor size, hard drive capacity) play important roles in describing bundle utilities. Model 6, which incorporates heterogeneity, increases the in-sample fit by 27% and hit rates by 13%, 17%, and 18%, respectively, compared with those of Model 5. Furthermore, Model 7, which incorporates two latent classes, increases the in-sample fit by 15% and hit rates by 17%, 15%, and 20%, respectively, compared with those of Model 6.

Reservation Prices

In our study, we asked each respondent to indicate his or her best bundle among all possible bundles by choosing one of three monitors, one of three PCs, and one of two printers described in Table 1 and how much he or she would be willing to pay for it. We use these data to test the validity of the estimated reservation prices.

Having determined Model 7 as the best for our data, we used it to calculate the reservation prices for the best bundle specified by each respondent. As mentioned previously, the reservation price of a bundle depends on which alternatives are included in the choice set. We therefore calculate the reservation prices of the best bundles in a choice set consisting of the best bundle and the no-buy option

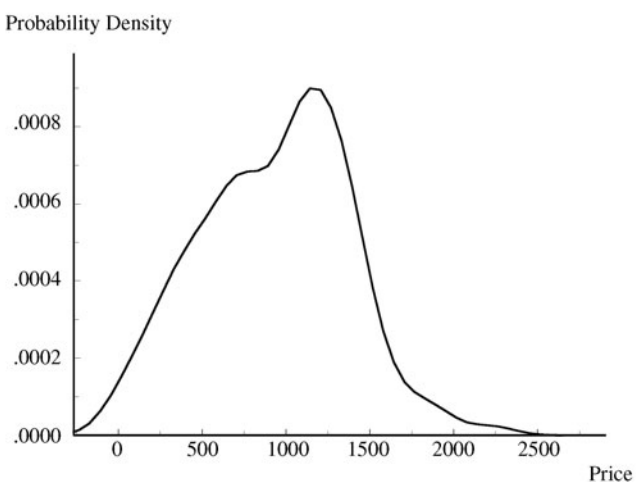
Table 7

IN-SAMPLE AND OUT-OF-SAMPLE FITS FOR ALTERNATIVE MODELS

Model	*In-Sample Fit (Log Marginal Likelihood)*	*Out-of-Sample Fit (Hit Rate)* 3	4	5
Model 1	–902.23	.55	.45	.38
Model 2	–1330.78	.41	.33	.28
Model 3	–1199.53	.46	.41	.34
Model 4	–1102.21	.50	.42	.36
Model 5	–1012.33	.58	.49	.42
Model 6	–820.55	.67	.59	.51
Model 7	–701.14	.74	.77	.61

Figure 3

INDIVIDUAL RESERVATION PRICES



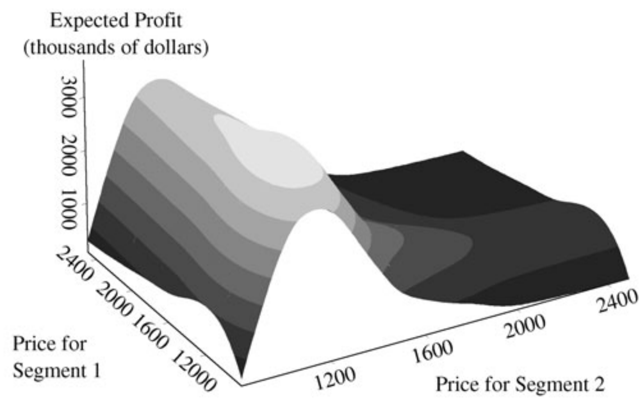
only.[13] Therefore, the reservation price of the best bundle in this choice set does not depend on the hypothetical bundles shown in our choice sets. We compare the estimated reservation price for the best bundle with the respondent-stated reservation price for that bundle and compute the mean absolute percentage error (MAPE). The values of MAPE ranged from a low of 17.01 for Model 7 to a high of 40.00 for Model 2. The lowest value of MAPE for Model 7 indicates that our model is useful for estimating reservation prices.

To illustrate the heterogeneity distribution of reservation prices for a bundle, we choose the first bundle out of the three bundles in a choice task (Bundle 1: ViewSonic 17″ monitor, eMachine eTower PC, and HP laser printer 1100; Bundle 2: Sony 21″ monitor, Dell PC, and HP laser printer; Bundle 3: ViewSonic 15″ monitor, HP PC, and Epson Inkjet) and obtained the distribution for estimated means of conditional individual reservation prices by using S-plus density estimation. This distribution, shown in Figure 3,

[13]In these calculations, we excluded 53 respondents who were asked to respond to some choice tasks before the reservation price question for their best bundles because their reservation prices could reflect the influences of competing brands shown in the choice tasks.

Figure 4

THE EFFECTS OF BUNDLE PRICES ON PROFITS



reflects considerable heterogeneity in reservation prices among respondents.

Optimal Price Analysis

Optimal prices depend on each firm's business goal, such as maximization of profit, revenue, or market share. We assume that the firm's goal is profit maximization, and the hypothetical cost for Bundle 1 is 70% of the price, \$714. The sample used in our study is assumed to be representative of the target market, and the size of the target market is assumed to be 1000 times as large as the sample size (projection factor M is equal to 1000). On that assumption, we calculate optimal prices for the bundle under the strategy of pure bundling.[14]

We calculate three optimal prices according to two marketing strategies: (1) optimal single price for the whole market (P1) and (2) two optimal prices for each segment market (P2). Our modeling approach enables us to calculate optimal prices for each consumer if necessary (under a 1–1 marketing strategy). For the strategy P1, the optimal single price for the bundle was \$1,197, with an expected profit of \$2,588,800. For strategy P2, the optimal price was \$1,310 for Segment 1 and \$1,125 for Segment 2. The expected profit based on strategy P2 is \$2,620,200. Figure 4 presents the relationship between segment-specific bundle prices and the corresponding expected profits.[15]

If the direct marketer (manufacturer or reseller) selling the bundle has the same cost structure as assumed in our analysis and wants to maximize profits, we conclude that the current price of \$1,020 for the bundle (obtained from the price lists) should be increased by \$177 to obtain the maximum expected profit of \$2,588,800. If the firm can sepa-

[14] We computed optimal prices by averaging optimal prices calculated with each parameter vector drawn at each iteration through MCMC simulation. We exclude 2.4% of total draws, in which price coefficients happened to be positive values, violating the regularity condition for the existence of optimal price as described in the following section.

[15]These optimal prices are somewhat lower than those estimated with the results for multinomial logit when the no-buy option is ignored. The actual estimates are \$1,230 for P1 and \$1,334 and \$1,214 for the two segments under P2. The estimated correlation for the two nests in nested multinomial logit is .91, which explains this difference.

rately access the market segments identified in our analysis and can implement a two-price strategy aimed at the two segments, our model suggests that the firm can increase the expected profit by $31,400 for a market size of 1000 potential customers.

SUMMARY AND FUTURE RESEARCH DIRECTIONS

Our research makes a contribution to the bundle literature in three ways. First, unlike other attribute-based models, which are developed only for assortments (bundles with homogeneous products), our model can be applied to any type of bundles, including bundles with heterogeneous products. In addition, our model helps researchers understand the relationship between two different approaches, the conjoint model and the balance model, because both of these models are special cases. Second, our model provides a general random utility model for bundle choices regardless of the type of bundles. Therefore, unlike the extant bundle models focusing on bundle pricing (Hanson and Martin 1990; Venkatesh and Mahajan 1993), our model enables researchers to study not only bundle pricing but also a variety of other problems that have not been studied in the bundle literature. For example, this is the first article to introduce customer segmentation for bundle markets based on interdependence among products with the help of a general utility model. Third, in addition to the advantages of attribute-based approach, our model has advantages against the extant models for optimal bundle pricing. The extant models for bundle pricing (Hanson and Martin 1990; Venkatesh and Mahajan 1993) use the reservation prices obtained from direct questioning as crucial input in estimating optimal prices for bundles. The validity of the optimal prices estimated from the extant models relies heavily on the accuracy of the reservation prices from direct questioning, and these direct estimates may be less reliable especially for infrequently purchased goods such as systems (Gabor and Granger 1965). Kalish and Nelson (1991) report that the predictive validity of reservation prices from direct questioning for choice behavior is significantly inferior to preference data. Several alternative methods to elicit reservation prices have been suggested in the economics and marketing literature (Wertenbroch and Skiera 2002), reflecting the difficulty of the elicitation of the true reservation price. In addition, the extant models for optimal bundle pricing do not fully consider individual reservation prices and may lead to suboptimal prices. For example, Hanson and Martin's (1990) model does not use individual reservation prices but reservation prices of customer segments as a proxy variable. If individual reservation price varies within a segment as shown in our study, the use of segment reservation price can lead to biased optimal price.

In contrast, instead of asking consumers for their reservation prices as Hanson and Martin (1990) do, our model uses consumers' stated choice data, which are empirically proved to be reliable and valid data for the prediction of consumer choice behavior for optimal bundle pricing. Therefore, the optimal prices estimated from our model are likely to be more valid than those obtained from the extant model using reservation prices. Furthermore, our model lessens respondents' burden and enables estimation of distinct reservation prices for each consumer in empirically determined customer segments.

Nevertheless, there are several limitations of our model as follows: First, our study does not consider a mixed bundling situation because the extension of the model for the mixed bundling situation is likely to detract from the core characteristics of the model without additional methodological novelty. We could extend our model to mixed bundling situations by suitably computing the sums and dispersion terms in the model. For example, the terms of dispersions will naturally be zero for the singletons in the set of options for the mixed bundling, and the terms for all other subsets will be similar to what we have used. However, as stated previously, optimal prices obtained from this model under pure bundling can be suboptimal because pure bundling may be a suboptimal strategy when components for bundles are not strongly complementary (Venkatesh and Kamakura 2003). A future research direction is to extend the model to mixed bundling.

Second, the classification of attributes according to comparability may not be the same for all consumers. A useful research direction is to develop suitable methods to measure comparability among consumers and new tests for homogeneity of classification among consumers. Third, some other types of bundle evaluation processes may exist, which can weaken consumers' comparison of bundle components for "balancing" behavior. For example, consumers might construe a strong brand in a bundle as a credibility bond for a weaker brand in the bundle.[16] Further research can explore such effects and develop a model to incorporate them. Such a model can also be used to study the effects of cobranding on consumer multicategory choices.

We identify two additional directions for further research. First, our model can be applied to determine the optimal bundle composition and its price to maximize profit. Here, the major research issue is the translation of the attribute levels to physical design features of the components in the bundle. Second, our model can be extended to dynamic multiple-category choice problems in which consumers purchase products over time.

APPENDIX: MCMC USING GIBBS SAMPLER FOR INFERENCES

We use priors and hyper priors as follows. Note that we employ data augmentation in estimating the model by introducing Z_i. Sampling a vector of preference parameters, ϕ, from a full-conditional distribution can be significantly simplified by introducing unobserved variables, Z_i, that indicate to which segment consumer i belongs (Albert and Chib 1993; Tanner and Wong 1987) as follows:

Prior

$$\rho_i = (\rho_{i1}, \rho_{i0})' \sim N_2(\mu_\rho, T_\rho),$$

where $\mu_\rho = (\mu_{\rho 1}, \mu_{\rho 0})'$ and T_ρ is 2×2 covariance matrix.

$$\left(\phi_i \middle| z_i = f\right) \sim N_R(\mu_f, T_f),$$

where $\phi_i = (\alpha_i, \beta_i, \gamma_i)'$, $\mu_f = (\mu_{f1}, \mu_{f2}, ..., \mu_{fR})'$, and T_f is $R \times R$ covariance matrix.

[16]We thank a reviewer for pointing this out.

$$Z_i \sim \text{Multinomial}(\psi_{1i}, \psi_{2i}, .., \psi_{Fi}, n = 1),$$

where $Z_i = f$ if consumer i belongs to segment f.

Hyper Prior

$$\mu_\rho \sim N_2(\upsilon_\rho, \Omega_\rho) \text{ with } \upsilon_\rho = 0 \times I_{2\times 2} \text{ and } \Omega_\rho = 10^4 I_{2\times 2};$$

$$\mu_f \sim N_R(\upsilon_f, \Omega_f) \text{ with } \upsilon_f = 0 \times I_{R\times R} \text{ and } \Omega_f = 10^4 I_{R\times R}$$

$$\text{for } f = 1, 2, ..., F;$$

$$T_\rho^{-1} \sim \text{Wishart}(v_\rho, S_\rho) \text{ with } v_\rho = 4 \text{ and } S_\rho = v_\rho I_{2\times 2};$$

$$T_f^{-1} \sim \text{Wishart}(v_f, S_f) \text{ with } v_f = R + 2 \text{ and } S_o = v_f I_{R\times R}$$

$$\text{for } f = 1, 2, ..., F; \text{ and}$$

$$\psi \sim \text{Ordered Dirichlet}(\omega_1, \omega_2, ..., \omega_F),$$

where $\psi = (\psi_1, \psi_2, ..., \psi_F)'$ and $\omega_1 < \omega_2 < ... < \omega_F$.

Conjugate families are specified for hyper priors (μ_ρ, μ_f, T_ρ, and T_f) with diffuse hyper-hyper priors (υ_ρ, Ω_ρ, v_ρ, S_ρ, υ_f, Ω_f, v_f, and S_f) that reflect the lack of information on hyper priors. We specify a Dirichlet distribution for ψ and impose an ordinal restriction on ω for model identification (Diebolt and Robert 1994; McLachlan and Basford 1988).

Posteriors

The posterior distributions are not determined in a closed form. They contain intractable integrals because nonconjugate families are specified for priors. We therefore use simulation-based inferences, MCMC, by drawing simulated samples of parameter values from posterior distributions using the Gibbs sampler (Albert and Chib 1993). We implement the sampling by alternating the generation of augmented data, Z, and all parameters, $\theta = (\rho, \phi, \mu, T, \omega, \psi)$, iteratively, starting from an initial vector, $Z^{(0)}$ and $\theta^{(0)} = (\rho^{(0)}, \phi^{(0)}, \mu^{(0)}, T^{(0)}, \omega^{(0)}, \psi^{(0)})$. Each parameter vector is drawn for each iteration, t, from its full-conditional distributions according to the following procedures until convergence.

The Augmented Data Generation

(A1) Generate $Z_i^{(t+1)} \sim \text{Multinomial}\left(\xi_{1i}^{(t+1)}, \xi_{2i}^{(t+1)}, ..., \xi_{Fi}^{(t+1)}, n = 1\right)$, where

$$\xi_{fi}^{(t+1)} = \left|T_f^{(t)}\right|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left(\phi_i^{(t)} - \mu_f^{(t)}\right)' T_f^{(t)-1} \left(\phi_i^{(t)} - \mu_f^{(t)}\right)\right]\psi_{fi} \Bigg/ \sum_{f'=1}^{F} \left|T_{f'}^{(t)}\right|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left(\phi_i^{(t)} - \mu_{f'}^{(t)}\right)' T_{f'}^{(t)-1} \left(\phi_i^{(t)} - \mu_{f'}^{(t)}\right)\right]\psi_{f'i}.$$

The Model Parameter Generation

Generate parameters $\theta^{(t+1)} = (\phi^{(t+1)}, \mu^{(t+1)}, T^{(t+1)}, \omega^{(t+1)}, \psi^{(t+1)})$ given augmented data $Z^{(t+1)}$.

(A2) Generate $\rho^{(t+1)}$.

$$\rho_i^{(t+1)} \propto \left\{\prod_{q=1}^{Q} P_{i0|B(q)}\left(\rho_i^{(t+1)}\right)^{D_{i|B(q)}} \left[1 - P_{i0|B(q)}\left(\rho_i^{(t+1)}\right)\right]^{(1-D_{i|B(q)})}\right\} \left\{\left|T_\rho^{(t)}\right|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left(\rho_i^{(t+1)} - \mu_\rho^{(t)}\right)'\left[T_\rho^{(t)}\right]^{-1}\left(\rho_i^{(t+1)} - \mu_\rho^{(t)}\right)\right]\right\},$$

$$\text{where } D_{i|B(q)} = \begin{cases} 1 & \text{if consumer i chooses the no-buy option in the qth choice set,} \\ 0 & \text{if consumer i chooses any bundle in the qth choice set.} \end{cases}$$

(A3) Generate $\phi^{(t+1)}$.

$$\phi_i^{(t+1)} \propto \left[\prod_{q=1}^{Q}\prod_{b\in B(q)} P_{ib|B(q)}\left(\phi_i^{(t+1)}\right)^{Y_{ib|B(q)}}\right] \left\{\left|T_f^{(t)}\right|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left(\phi_i^{(t+1)} - \mu_f^{(t)}\right)'\left[T_f^{(t)}\right]^{-1}\left(\phi_i^{(t+1)} - \mu_f^{(t)}\right)\right]\right\},$$

$$\text{if } Z_i^{(t+1)} = f.$$

We implement sampling for ρ and ϕ through a random-walk Metropolis–Hastings algorithm using normal distribution with zero mean and some standard deviation adjusted at each iteration, because nonconjugate families are specified for their priors. (For details, see Robert and Casella 1999.) We used different values for the standard deviation, varying from .3 to 2, at each iteration to generate well-mixed draws (the average acceptance rate was 3.54).

(A4) Generate $T_f^{(t+1)}$ for $f = 1, 2, ..., F$.

$$T_f^{-1(t+1)} \sim \text{Wishart}_R\left[v_o + N_f^{(t+1)}, S_o + \sum_{i=1}^{I}\left(\phi_i^{(t+1)} - \mu_f^{(t+1)}\right)\left(\phi_i^{(t+1)} - \mu_f^{(t+1)}\right)' 1\left(Z_i^{(t+1)} = f\right)\right],$$

$$\text{where } N_f^{(t+1)} = \sum_{i=1}^{I} 1\left(Z_i^{(t+1)} = f\right).$$

(A5) Generate $\mu_f^{(t+1)} = \left(\mu_{f1}^{(t+1)}, \mu_{f2}^{(t+1)}, ..., \mu_{fR}^{(t+1)}\right)'$ for $f = 1, 2, ..., F$.

$$\mu_f^{(t+1)} \sim N_R\left[\left(N_f^{(t+1)}T_f^{-1(t+1)} + \Omega^{-1}\right)^{-1} N_f^{(t+1)}T_f^{-1(t+1)}\bar{\phi}_f^{(t+1)}, \left(N_f^{(t+1)}T_f^{-1(t+1)} + \Omega^{-1}\right)^{-1}\right],$$

$$\text{where } \bar{\phi}_f^{(t+1)} = \frac{\sum_{i=1}^{I}\phi_i^{(t+1)} 1\left(Z_i^{(t+1)} = f\right)}{N_f^{(t+1)}} \text{ and } \Omega = 10^4 I_{R\times R}.$$

(A6) Generate $\omega_f^{(t+1)}$ for all f.

$$\omega_f^{(t+1)} \sim \text{Gamma}\left(N_f^{(t+1)} + 1, 1\right) \text{ with}$$
$$\omega_1^{(t+1)} < \omega_2^{(t+1)} < \ldots < \omega_F^{(t+1)}.$$

(A7) Calculate $\psi_f^{(t+1)} = \dfrac{\omega_f^{(t+1)}}{\sum_{f'=1}^{F} \omega_{f'}^{(t+1)}}$ for all f.

When these iterations are stopped after convergence, estimates of parameters of interest are obtained from samples. In our study, three Markov chains are generated, in which 15,000 draws are implemented in total for the burn-in period with convergence tests, and 3000 draws (1000 draws from three chains, respectively) are used for inference.

REFERENCES

Ansari, Asim, S. Siddarth, and Charles B. Weinberg (1996), "Pricing a Bundle of Products and Services: The Case of Nonprofits," *Journal of Marketing Research*, 33 (February), 86–93.

Albert, James and Siddhartha Chib (1993), "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of the American Statistical Association*, 88 (422), 669–79.

Allenby, Greg M., Neeraj Arora, and James L. Ginter (1998), "On the Heterogeneity of Demand," *Journal of Marketing Research*, 35 (August), 384–89.

Bradlow, Eric and Vithala R. Rao (2000), "A Hierarchical Bayes Model for Assortment Choice," *Journal of Marketing Research*, 37 (May), 259–68.

Chib, Siddhartha and Ivan Jeliazkov (2001), "Marginal Likelihood from the Metropolis–Hastings Output," *Journal of the American Statistical Association*, 96 (432), 270–81.

Chickering, D.M. and D. Heckerman (1997), "Efficient Approximations for the Marginal Likelihood of Bayesian Networks with Hidden Variables," *Machine Learning*, 29 (2–3), 181–212.

Diebolt, Jean and Christian P. Robert (1994), "Estimation of Finite Mixture Distributions Through Bayesian Sampling," *Journal of the Royal Statistical Society*, Series B (Methodological), 56 (2), 363–75.

Farquhar, Peter H. and Vithala R. Rao (1976), "A Balance Model for Evaluating Subsets of Multiattributed Items," *Management Science*, 22 (January), 528–39.

Gabor, Andre and Clive W.J. Granger (1965), "The Pricing of New Products," *Scientific Business*, 10 (August), 141–50.

Goldberg, Stephen M., Paul E. Green, and Yoram Wind (1984), "Conjoint Analysis of Price Premiums for Hotel Amenities," *Journal of Business*, 54 (Supplement 1, Part 2), S111–32.

Green, Paul E. and Michael T. Devita (1974), "A Complementarity Model of Consumer Utility for Item Collections," *Journal of Consumer Research*, 1 (December), 56–67.

———, Yoram Wind, and Arun K. Jain (1972), "Preference Measurement of Item Collections," *Journal of Marketing Research*, 9 (November), 371–77.

Hanson, Ward A. and R. Kipp Martin (1990), "Optimal Bundle Pricing," *Management Science*, 36 (February), 155–74.

Harlam, Bari A. and Leonard M. Lodish (1995), "Modeling Consumers' Choices of Multiple Items," *Journal of Marketing Research*, 32 (November), 404–18.

Johnson, Michael D. (1984), "Consumer Choice Strategies for Comparing Noncomparable Alternatives," *Journal of Consumer Research*, 11 (December), 741–53.

Kalish, Shlomo and Paul Nelson (1991), "A Comparison of Ranking, Rating, and Reservation Price Measurement in Conjoint Analysis," *Marketing Letters*, 2 (4), 327–35.

Kamakura, Wagner A. and Gary J. Russell (1989), "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure," *Journal of Marketing Research*, 26 (November) 379–90.

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

Lattin, James M. (1987), "A Model of Balanced Choice Behavior," *Marketing Science*, 6 (Winter), 48–65.

Louviere, Jordan J., David A. Hensher, and Joffre D. Swait (2000), *Stated Choice Methods: Analysis and Application*. Cambridge, U.K.; New York: Cambridge University Press.

Manchanda, Puneet, Asim Ansari, and Sunil Gupta (1999), "The 'Shopping Basket': A Model for Multi-category Purchase Incidence Decisions," *Marketing Science*, 18 (2), 95–115.

McAlister, Leigh (1979), "Choosing Multiple Items from a Product Class," *Journal of Consumer Research*, 6 (December), 213–24.

——— (1982), "A Dynamic Attribute Satiation Model of Variety Seeking Behavior," *Journal of Consumer Research*, 9 (September), 141–50.

McFadden, Daniel L. (1981), "Econometric Models of Probabilistic Choice," in *Structural Analysis of Discrete Data with Econometric Applications*, C.F. Manski and D. McFadden, eds. Cambridge, MA: MIT Press, 198–272.

McLachlan, Geoffrey J. and Kaye E. Basford (1988), *Mixture Models: Inference and Applications to Clustering*. New York: Marcel Dekker.

Meliǎ, Marina and David Heckerman (1998), "An Experimental Comparison of Several Clustering and Initialization Methods," *Machine Learning*, 42 (3), 9–29.

Newton, Michael A. and Adrian E. Raftery (1994), "Approximate Bayesian Inference with the Weighted Likelihood Bootstrap," *Journal of the Royal Statistical Society*, Series B (Methodological), 56 (1), 3–48.

Poirier, Dale J. (1996), "A Bayesian Analysis of Nested Logit Models," *Journal of Econometrics*, 75 (November), 163–81.

Robert, Christine P. and George Casella (1999), *Monte Carlo Statistical Methods*. New York: Springer-Verlag.

Russell, Gary J., S. Ratneshwar, Allan D. Shocker, et al. (1999), "Multiple-Category Decision-Making: Review and Synthesis," *Marketing Letters*, 10 (3), 319–32.

——— and Ann Petersen (2000), "Analysis of Cross-Category Dependence in Market Basket Selection," *Journal of Retailing*, 76 (Fall), 367–92.

Schmalensee, R. (1984), "Gaussian Demand and Commodity Bundling," *Journal of Business*, 57 (January, Part 2), S221–S230.

Tanner, Martin A. and Wing H. Wong (1987), "The Calculation of Posterior Distributions by Data Augmentation," *Journal of the American Statistical Association*, 82 (398), 528–40.

Venkatesh, R. and Wagner A. Kamakura (2003), "Optimal Bundling and Pricing Under a Monopoly: Contrasting Complements and Substitutes from Independently Valued Products," *Journal of Business*, forthcoming.

——— and Vijay Mahajan (1993), "A Probabilistic Approach to Pricing a Bundle of Prices or Services," *Journal of Marketing Research*, 30 (November), 494–508.

——— and ——— (1997), "Products with Branded Components: An Approach for Premium Pricing and Partner Selection," *Marketing Science*, 16 (2), 146–65.

Wertenbroch, Klaus and Bernd Skiera (2002), "Measuring Consumer Willingness to Pay at the Point of Purchase," *Journal of Marketing Research*, 31 (May), 228–41.

Wilson, Lynn O., Allen M. Weiss, and George John (1990), "Unbundling of Industrial Systems," *Journal of Marketing Research*, 27 (May), 123–38.

Wuebker, Georg and Vijay Mahajan (1999), "A Conjoint Analysis-Based Procedure to Measure Reservation Price and Optimally Price Product Bundles," in *Optimal Bundling: Marketing Strategies for Improving Economic Performance*, Ralph Fuerderer, Andreas Herrmann, Georg Wuebker, and H. Hennig-Schmidt, eds. New York: Springer-Verlag, 157–74.