Measuring Heterogeneous Reservation Prices for Product Bundles

Kamel Jedidi • Sharan Jagpal • Puneet Manchanda
Graduate School of Business, Columbia University, New York 10027
Rutgers Business School, Rutgers University, 180 University Avenue, Newark, New Jersey 07102
Graduate School of Business, University of Chicago, 1101 East 58th Street, Chicago, Illinois 60637
kj7@columbia.edu • jagpal@andromeda.rutgers.edu • puneet.manchanda@gsb.uchicago.edu

This paper develops a model for capturing continuous heterogeneity in the joint distribution of reservation prices for products and bundles. Our model is derived from utility theory and captures both within- and among-subject variability. Furthermore, it provides dollarmetric reservation prices and individual-level estimates that allow the firm to target customers and develop customized and nonlinear pricing policies.

Our experiments show that, regardless of whether the products are durables or nondurables, the model captures heterogeneity and predicts well. Models that assume homogeneity perform poorly, especially in predicting choice of the bundle. Furthermore, the methodology is robust even when respondents evaluate few profiles.

Self-stated reservation prices do not have any informational content beyond that contained in the basic model. The direct elicitation method appears to understate (overstate) the variation in reservation prices across consumers for low-priced (high-priced) products and bundles. Hence this method yields biased demand estimates and leads to suboptimal product-line pricing policy.

The optimization results show that the product-line pricing policy depends on the degree of heterogeneity in the reservation prices of the individual products and the bundle. A uniformly high-price strategy for all products and bundles is optimal when heterogeneity is high. Otherwise, a hybrid strategy is optimal.

(*Bundling; Reservation Prices; Optimal Pricing; Consumer Heterogeneity; Multinomial Probit Models; Conjoint Analysis; Bayesian Models*)

1. Introduction

Bundling is prevalent across a wide spectrum of industries ranging from consumer durable products (e.g., automobile options) to high-technology products (e.g., software suites such as Microsoft's Office 2000) to the service sector (including nonprofit organizations). Furthermore, bundling strategies are becoming increasingly important as strategic partnerships and strategic alliances proliferate, especially among firms that sell information goods on the Internet.

Although the theoretical literature on bundling has burgeoned (see §2 for a detailed review), three important methodological questions remain: Can one develop a general empirical methodology for capturing heterogeneity in reservation prices? Can the methodology capture general patterns of substitutability and complementarity? Is the method robust for durables and nondurables?

This paper has three main objectives. The first is to develop and test a theory-based model that captures continuous heterogeneity in the joint distribution of reservation prices for individual products and bundles. The second is to demonstrate how one can test theoretical assumptions (e.g., complementarity

0732-2399/03/2201/0107$5.00
1526-548X electronic ISSN

and substitutability) using the model. The third is to develop and test a flexible algorithm that allows the practitioner to use the model estimates to choose the optimal product-line pricing policy.

Methodologically, our model is derived from utility theory and allows one to capture both heterogeneity and consumer error in judgment. Hence the model can be used to capture general patterns of substitutability and complementarity.

Our method provides dollarmetric reservation prices and allows the user to obtain individual-level estimates. Hence, the manager can use the results to choose optimal product-line pricing policy; in particular, he/she can develop customized and nonlinear pricing policies. In addition, policy makers can use the model to assess the social welfare implications of bundling pricing policies in different industries (see Ansari et al. 1996). As discussed in §3, the method is easy to implement and is computationally efficient.

We conduct two experimental studies to examine model robustness. To assess the generalizability of the model, each study examines three product groups containing both durable and nondurable products. The nondurable group includes six-month subscriptions to *Time* and *Newsweek* magazines. The two durable groups were, respectively, a video camera/video cassette recorder and a microwave oven/television.

The first study has two main objectives: to test whether it is necessary for the model to include heterogeneity and to demonstrate how one can use the method to test theoretical assumptions. To increase the validity of the experiment and to improve statistical efficiency, the experimental design uses a hybrid Latin-square design with within-cell randomization.

The results show that for all three product groups, our model performs well based on both fit and predictive accuracy. Specifically, the model predicts consumers' choices reasonably well. Thus, the model can be used for targeting consumers. In contrast, models that assume homogeneity perform poorly, especially in identifying consumers who will purchase the bundle. The results also show that bundling reduces consumer heterogeneity. However, in general, individual-level reservation prices are neither additive nor statistically independent.

The second study further tests the robustness of the method for both durables and nondurables. The main objectives are threefold. First, we test whether model performance deteriorates significantly if respondents only evaluate a limited number of profiles. Second, we compare our model with the direct elicitation method, in which respondents explicitly state their reservation prices for the individual products and the bundle. Third, we test if model performance can be improved by using an augmented model that includes self-stated reservation prices as an additional regressor.

The results show that our model captures heterogeneity reasonably well, even when respondents evaluate a limited number of profiles (three per product group in the second study and nine per product group in the first). Regardless of product group (durable or nondurable), our model significantly outperforms the direct elicitation method. Specifically, the direct elicitation method appears to understate (overstate) the degree of heterogeneity in the reservation prices for low-priced (high-priced) products. Consequently, this method yields biased estimates of the appropriate demand curves and leads to suboptimal product-line pricing policy. In general, the augmented model does not improve fit significantly. That is, self-stated reservation prices do not provide additional information beyond the basic heterogeneity model.

The optimization results consistently show that the firm should use a mixed bundling strategy. However, the optimal product-line policy depends on the degree of heterogeneity in the reservation prices of the individual products and the bundle. For example, for the *Time-Newsweek* group, a uniformly high-price strategy is optimal (i.e., the prices for the individual products and the bundle exceed the appropriate mean reservation prices). In contrast, for the videocassette player/video camera group, a hybrid strategy is optimal in which the firm uses a high-price approach for the videocassette player but prices the video camera and the bundle below the corresponding mean reservation prices.

The optimization results also show that the direct elicitation method leads to a suboptimal pricing strategy. For two of the three product groups, the self-stated method provided significantly inflated measures of heterogeneity for the individual products and bundles. Hence, this method led to excessively high "optimal" prices for the individual products and the bundle. However, for the third group, the direction of bias in the heterogeneity of reservation prices varied across products and the bundle. Consequently, the effect of the self-stated method on product-line prices cannot be specified a priori.

The structure of the paper is as follows. Section 2 reviews the theoretical and empirical research on bundling. Section 3 describes our methodology. Section 4 discusses the experimental design and data collection procedure. Section 5 analyzes the empirical results. Section 6 discusses the optimization results when the model estimates are used to choose the product-line pricing policy. Section 7 discusses model extensions. Section 8 summarizes the main results and suggests directions for future research.

2. Literature Review

This section reviews the extant theoretical and empirical research on bundling.[1]

2.1. Theoretical Research on Bundling

Bundling research has focused on the monopoly and duopoly cases. We first review monopoly models.

Adams and Yellen (1976) developed a two-product, monopoly bundling model using two key assumptions: The reservation prices for products are additive and negatively correlated. Using numerical examples, they demonstrated that, in general, the firm should use a mixed bundling strategy (i.e., consumers should be allowed to choose from the individual products and the bundle). Lewbel (1985) generalized the Adams-Yellen framework to allow for substitutes and complements. His research shows that, depending on the degree of substitutability/complementarity, the monopolist may choose a pure bundling strategy or even choose not to offer a bundle at all.

Schmalansee (1984) developed a two-product monopoly bundling model in which he relaxed the assumption that the reservation prices of the individual products are negatively correlated. However, he retained the additivity assumption. Schmalansee's main results are twofold. First, bundling can increase profits even if the reservation prices of the individual products in a product line are positively correlated. Second, the optimum product-line policy depends on the standardized markups for the products and bundles (i.e., the ratios of the unit profit markups on each product/bundle to the standard deviations of the reservation prices for the corresponding product/bundle).

Subsequent theoretical research has extended Schmalansee's monopoly model. McAfee et al. (1989) and Salinger (1995) deal with the two-product case, and Bakos and Brynjolfsson (2000) deal with the general case where the bundle contains more than two products. All these studies assume that the reservation prices of the individual products in the bundle are additive, statistically independent, or both. Venkatesh and Kamakura (2003) relax the additivity assumption to allow for certain patterns of substitutability and complementarity (see θ in Equation (2) of their paper); however, it retains the assumption of statistical independence. Based on simulation evidence, Venkatesh and Kamakura conclude that the monopolist's optimal bundling strategy depends on the marginal costs of producing the products (bundles) and the degree of substitutability and complementarity.

Recent research has focused on the duopoly case where firms produce complementary products. Economides (1993) concludes that the subgame perfect Nash equilibrium is for each duopolist to offer a mixed bundle. In contrast, Anderson and Leruth (1993) conclude that the optimal duopoly strategy is to sell each product separately. Kopalle et al. (1999) attempt to reconcile these results using a market-growth model. Specifically, they show that, as the scope for market expansion decreases, the equilibrium bundling strategy shifts from mixed bundling to selling pure components only.

[1] We define bundling as the joint sale of several products containing different attributes. This definition differs from that used in several marketing papers (e.g., Villas-Boas 1998 and Desai 2001). Those papers define bundling as the separate sale of products in a product assortment, where all products contain the same attribute(s).

2.2. Empirical Research on Bundling

Normative empirical bundling studies have typically followed one of two approaches: Either use self-stated reservation prices as a proxy for the unobservable true reservation prices (Hanson and Martin 1990, Venkatesh and Mahajan 1993, Ansari et al. 1996) or estimate bundle preferences using conjoint-based experiments (Goldberg et al. 1984, Wuebecker and Mahajan 1999).[2] An alternative approach, not used in the bundling literature, is to use an auction-based method (Vickrey 1961, Becker et al. 1964, Kagel 1995). Following is a brief review of these methods.

The use of self-stated reservation prices makes two implicit assumptions: reservation prices are deterministic (i.e., within-subject variability is zero) and provide unbiased estimates of the true, unobservable reservation prices. However, as pointed out by Gabor and Granger (1965), self-stated reservation prices can lead to significant measurement error, especially for infrequently purchased products. In addition, as Monroe (1990, p. 107–112) notes, self-stated reservation prices are likely to be biased downward because consumers feel that it is in their best interest to keep prices down. Thus, unless measurement error is small, the use of self-stated reservation prices is likely to lead to biased estimates of the demand curves for the individual products and the bundle, resulting in suboptimal product-line pricing policies.

An alternative strategy is to use conjoint-type experiments. Goldberg et al. (1984) (henceforth GGW) used categorical conjoint analysis (canonical correlation) to measure consumers' preferences for bundles. The GGW method, however, does not provide estimates of reservation prices or allow for heterogeneity (the focus of our paper). Recently, Wuebecker and Mahajan (1999) developed a conjoint-based approach to estimate individual-level reservation prices for the bundle only. This method, however, assumes statistical independence and does not provide estimates of the reservation prices of the individual products comprising the bundle.

[2] Noble and Gruca (1999) recently conducted a descriptive empirical study on the use of different pricing methods in industrial markets. They found that bundle pricing was more common in highly price-elastic markets.

Auction-based methods, in contrast to conjoint experiments, measure real and not self-stated choices. However, they are not a panacea. Auction-based estimates are likely to be biased upward because respondents bid amounts that are higher than their reservation prices (Kagel 1995, Kagel et al. 1987). In addition, auction methods make several restrictive assumptions that could seriously compromise internal and external validity: (a) Reservation prices are deterministic (an assumption likely to be violated for new products or products with which the respondent is unfamiliar); (b) reservation prices are independent of market prices (this assumption precludes consumers from using price as a signal of quality); and (c) the price-setting mechanism in the market is random and unknown to subjects. Finally, these methods cannot be used to determine reservation prices for durables (Wertenbroch and Skiera 2002).

In summary, the bundling literature shows that the firm's policy depends crucially on the joint distribution of reservation prices for products and bundles in the population. Hence, we need an empirical methodology that can capture continuous consumer heterogeneity and general forms of substitutability and complementarity. For generality, the method should be applicable to both durables and nondurables, allow for errors in consumer judgment, and provide dollarmetric measures of reservation prices. Our study focuses on these problems.

3. Method

3.1. Model

Suppose the firm's product line consists of q products, where $q \geq 2$. The firm's decision problem is then to choose an unbundled, pure, or mixed bundling strategy for these products. The consumer's decision problem is to choose the most preferred alternative from the available set of individual products and bundles. Though our model is general, for ease of exposition, we present the methodology for the $q = 2$ case. Hence, given the prices of the two products and the bundle, consumers choose from a menu of four alternatives: product 1 only, product 2 only, a bundle consisting of products 1 and 2, or no purchase.

To infer reservation prices, we use a choice-based experiment where we manipulate the prices of the choice alternatives and observe the choice outcomes. Let L_m denote the number of pricing levels for product/bundle m. Let P_{ml_m} be the lth price level of product/bundle m, where $l_m = 1, 2, \ldots, L_m$; $m = 1$ represents product 1; $m = 2$ represents product 2; and $m = 3$ represents the bundle. Thus, $(P_{1l_1}, P_{2l_2}, P_{3l_3})_j$ represents the jth pricing scenario for product 1, product 2 and the bundle, respectively, where $j = 1, \ldots, J$ and $J \le L_1 \times L_2 \times L_3$.

Assume that consumers are utility maximizers and have quasilinear (i.e., additively separable) utility functions.[3] Then, maximizing utility is equivalent to maximizing surplus (see the Appendix). In other words, for any given pricing scenario, each consumer i compares the price of each alternative (P_{ml_m}) with his/her reservation price for that alternative (R_{im}). Suppose the price of each alternative exceeds the reservation price for that alternative (i.e., $P_{ml_m} > R_{im}$ for all m). Then consumer i will choose the no-purchase alternative. Suppose the consumer surplus, $S_{iml_m} = R_{im} - P_{ml_m}$, is greater than zero for at least one alternative. Then, consumer i will choose the alternative m for which S_{iml_m} is the highest.

The no-purchase option is crucial. Including the no-purchase option provides four important advantages. First, including the no-purchase option in the choice set along with randomized prices across subjects eliminates specification error. Second, inclusion of the no-purchase option allows us to capture competitive effects. Third, this approach allows us to capture heterogeneity in price expectations and reference pricing effects. Fourth, as explained later (see §3), by including no purchase, the model allows us to obtain dollarmetric reservation prices. As discussed previously, this property is essential for choosing nonlinear pricing policies and measuring welfare.

[3] The Appendix explains the utility theory foundation of our model. Our quasilinear utility function assumes that for any given individual, the marginal utility of income is constant across products and bundles (see Appendix, Equation (A.1)). Clearly, this assumption may not hold if the product (bundle) prices are a significant proportion of the consumer's income.

To illustrate how the no-purchase option allows us to capture competitive effects, consider any price scenario for the products and bundles included in the experiment. Suppose the consumer chooses the no-purchase option. Then the consumer's choice implies that at least one of the following alternatives provides a greater surplus than the alternatives included in the experiment: no purchase at all, purchase of another product or bundle not included in the experiment (e.g., competitive offerings), or purchase of one of the alternatives in the experiment from another source (possibly at a cheaper price). Thus, competitive products are subsumed under the "no-purchase" option. In particular, the results will not be affected by the choice set (i.e., the set of products and bundles that are excluded from the experiment).[4]

To illustrate how the no-purchase option allows us to capture heterogeneous price expectations and differential reference price effects, consider a behavioral model in which the consumer derives two types of benefits from purchase: acquisition utility (the benefit from consuming the product/bundle) and transaction utility (a gain or loss that is based on the difference between the actual product/bundle price and the consumer's reference price). See Thaler (1986).

Suppose two consumers, A and B, have the same acquisition utilities for a Sony VCR but have heterogeneous price expectations (reference prices). Assume that in the experiment both A and B are given the following scenario: They can purchase the Sony VCR for \$240. But A expects to buy the Sony VCR for \$125 in the marketplace; in contrast, B believes the market price for the Sony VCR is \$425. Given this price scenario, A obtains a negative surplus (loss) of \$115 from purchasing the Sony VCR. However, B obtains a net surplus of \$185. Thus A chooses the no-purchase option in the experiment, whereas B chooses to purchase the Sony VCR. Our model captures both these behaviors. The reason is twofold. First, the model

[4] Suppose the choice set includes only products and bundles sold by the firm. Then the model will capture the joint distribution of reservation prices for these products and bundles. However, it will not measure the effects of the firm's policy on competitors. Suppose the firm seeks to measure such competitive effects. Then the only change necessary is to include the relevant competitive products and bundles in the experiment.

assumes that consumers maximize their net surplus. Second, the model explicitly incorporates heterogeneity. That is, it allows net surpluses to vary across individuals and products.

The choice experiment results in a binary indicator (choice/no choice) variable for each alternative for each consumer and price scenario. Therefore, let us define an indicator vector, $\mathbf{y}_i^j$, with elements y_{im}^j, $m = (1, 2, 3)$, for each individual i and price scenario $(P_{1l_1}, P_{2l_2}, P_{3l_3})_j$. Thus, if consumer i chooses alternative m given price scenario $j(j = 1, \ldots, J_i)$, then $y_{im}^j = 1$, otherwise $y_{im}^j = 0$. For example, $\mathbf{y}_i^j = (0, 0, 0)$ implies that consumer i chose the no-purchase alternative given price scenario j. Similarly, $y_i^j = (0, 0, 1)$ implies that consumer i chose the bundle given price scenario j.

Suppose individual i faces price scenario $(P_{1l_1}, P_{2l_2}, P_{3l_3})_j$. As in random utility models, define the latent reservation price for alternative m as follows:

$$R_{imj} = r_{im} + u_{imj}, \quad m = 1, 2, 3; \qquad i = 1, \ldots, I;\ j = 1, \ldots, J_i, \tag{1}$$

where r_{im} is the "true"[5] reservation price of consumer i for alternative m and u_{imj} is a random error that captures consumer errors in judgment. We assume $\mathbf{u}_{ij} = (u_{i1j}, u_{i2j}, u_{i3j})'$ to be distributed i.i.d. as multivariate normal $N(\mathbf{0}, \boldsymbol{\Sigma})$ where $\mathbf{0}$ denotes the null vector and $\boldsymbol{\Sigma}$ is the covariance matrix. Because Equation (1) is defined at the individual level, $\boldsymbol{\Sigma}$ captures the *within*-individual variability in reservation prices.[6]

Following the tradition in the hierarchical Bayesian models, we assume that the vector of true reservation prices, $\mathbf{r}_i = (r_{i1}, r_{i2}, r_{i3})'$, is distributed across the population of consumers as multivariate normal $N(\bar{\mathbf{r}}, \boldsymbol{\Lambda})$ where $\bar{\mathbf{r}} = (\bar{r}_1, \bar{r}_2, \bar{r}_3)'$ is the population mean reservation price vector and $\boldsymbol{\Lambda}$ denotes the covariance matrix of the "true" reservation prices across individuals. As Schmalansee (1984, p. S214) notes, "…the Gaussian [normal] distribution… seems much more plausible in this context [i.e., bundling] than the uniform distribution, for instance, that one would need to posit in order to justify more familiar looking linear demand curves." Thus, large diagonal elements of $\boldsymbol{\Lambda}$ denote high consumer heterogeneity in reservation prices. In contrast, large diagonal elements of $\boldsymbol{\Sigma}$ indicate large consumer errors in judgement for the appropriate products/bundle.

Note that the model does not impose any structure on the reservation prices. Specifically, the model does not assume that reservation prices are statistically independent since $\boldsymbol{\Delta}$ and $\boldsymbol{\Sigma}$ are nondiagonal. More importantly, as r_{i3} is not constrained, the bundle reservation price can be *superadditive* (i.e., $r_{i3} > r_{i1} + r_{i2}$), *additive* (i.e., $r_{i3} = r_{i1} + r_{i2}$), or *subadditive* (i.e., $r_{i3} < r_{i1} + r_{i2}$). Consequently, the model can be used to measure substitutability and complementarity and to test various bundling theories.

Given this structure, the conditional probability that consumer i chooses the no-purchase alternative under price scenario $(P_{1l_1}, P_{2l_2}, P_{3l_3})_j$ is

$$\begin{aligned}
&\Pr\{\mathbf{y}_i^j = (0, 0, 0) \mid \mathbf{r}_i\} \\
&\quad = \Pr\{R_{i1j} \le P_{1l_1}, R_{i2j} \le P_{2l_2}, R_{i3j} \le P_{3l_3}\} \\
&\quad = \Pr\{u_{i1j} \le P_{1l_1} - r_{i1}, u_{i2j} \le P_{2l_2} - r_{i2}, u_{i3j} \le P_{3l_3} - r_{i3}\} \\
&\quad = \int_{-\infty}^{P_{1l_1} - r_{i1}} \int_{-\infty}^{P_{2l_2} - r_{i2}} \\
&\qquad \cdot \int_{-\infty}^{P_{3l_3} - r_{i3}} f(u_{i1j}, u_{i2j}, u_{i3j})\, du_1\, du_2\, du_3,
\end{aligned} \tag{2}$$

where $f(u_{i1j}, u_{i2j}, u_{i3j})$ is the density function of error vector $\mathbf{u}_{ij}$ as described above.

The conditional probability of choosing product 1 is

$$\begin{aligned}
&\Pr\{\mathbf{y}_i^j = (1, 0, 0) \mid \mathbf{r}_i\} \\
&\quad = \Pr\{R_{i1j} \ge P_{1l_1}, (R_{i1j} - P_{1l_1}) \ge (R_{i2j} - P_{2l_2}), \\
&\qquad (R_{i1j} - P_{1l_1}) \ge (R_{i3j} - P_{2l_3})\} \\
&\quad = \int_{A_{11}}^{\infty} \int_{A_{12}}^{\infty} \int_{A_{13}}^{\infty} f(u_{i1j}, z_{i2j}^1, z_{i3j}^1)\, du_1\, dz_2^1\, dz_3^1,
\end{aligned} \tag{3}$$

where $A_{11} = P_{1l_1} - r_{i1}$, $A_{12} = (P_{1l_1} - r_{i1}) - (P_{2l_2} - r_{i2})$, $A_{13} = (P_{1l_1} - r_{i1}) - (P_{3l_3} - r_{i3})$, $z_{i2j}^1 = u_{i1j} - u_{i2j}$ and $z_{i3j}^1 = u_{i1j} - u_{i3j}$.

[5] We use the word "true" as in reliability theory where Observed Score = True Score + Error. In this sense, r_{im} could be interpreted as the mean reservation price for a given individual over repeated measurements.

[6] This formulation is consistent with Thurstone's utility model. That is, a consumer's response to the price profile for a given alternative is based on an instantaneous draw from that consumer's distribution of reservation prices for the alternative being evaluated (see Anderson et al. 1992, pp. 30–31).

Similarly, the conditional probability of choosing product 2 is

$$\begin{aligned}\Pr\{\mathbf{y}_i^j = (0,1,0) \mid \mathbf{r}_i\} &= \Pr\{R_{i2j} \geq P_{2l_2}, (R_{i2j} - P_{2l_2}) \geq (R_{i1j} - P_{1l_1}) \\ &\qquad (R_{i2j} - P_{2l_2}) \geq (R_{i3j} - P_{3l_3})\} \\ &= \int_{A_{22}}^{\infty}\int_{A_{21}}^{\infty}\int_{A_{23}}^{\infty} f(u_{i2j}, z_{i1j}^2, z_{i3j}^2)\, du_2\, dz_1^2\, dz_3^2, \end{aligned} \tag{4}$$

where $A_{22} = P_{2l_2} - r_{i2}$, $A_{21} = (P_{2l_2} - r_{i2}) - (P_{1l_1} - r_{i1})$, $A_{23} = (P_{2l_2} - r_{i2}) - (P_{3l_3} - r_{i3})$, $z_{i1j}^2 = u_{i2j} - u_{i1j}$, and $z_{i3j}^2 = u_{i2j} - u_{i3j}$.

Finally, the probability of choosing the bundle is

$$\begin{aligned}\Pr\{\mathbf{y}_i^j = (0,0,1) \mid \mathbf{r}_i\} &= \Pr\{R_{i3j} \geq P_{3l_3}, (R_{i3j} - P_{3l_3}) \geq (R_{i1j} - P_{1l_1}), \\ &\qquad (R_{i3j} - P_{3l_3}) \geq (R_{i2j} - P_{2l_2})\} \\ &= \int_{A_{33}}^{\infty}\int_{A_{31}}^{\infty}\int_{A_{32}}^{\infty} f(u_{i3j}, z_{i1j}^3, z_{i2j}^3)\, du_3\, dz_1^3\, dz_2^3, \end{aligned} \tag{5}$$

where $A_{33} = P_{3l_3} - r_{i3}$, $A_{31} = (P_{3l_3} - r_{i3}) - (P_{1l_1} - r_{i1})$, $A_{32} = (P_{3l_3} - r_{i3}) - (P_{2l_2} - r_{i2})$, $z_{i1j}^3 = u_{i3j} - u_{i1j}$, and $z_{i2j}^3 = u_{i3j} - u_{i2j}$.

Thus, the conditional likelihood, $\mathscr{L}_i \mid \mathbf{r}_i$, of observing the choices of consumer i selecting among the four choice alternatives under J_i pricing scenarios is

$$\begin{aligned}\mathscr{L}_i \mid \mathbf{r}_i = \prod_{j=1}^{J_i} &\Pr\{\mathbf{y}_i^j = (0,0,0)\}^{1-(y_{i1}^j + y_{i2}^j + y_{i3}^j)} \Pr\{\mathbf{y}_i^j = (1,0,0)\}^{y_{i1}^j} \\ &\cdot \Pr\{\mathbf{y}_i^j = (0,1,0)\}^{y_{i2}^j} \Pr\{\mathbf{y}_i^j = (0,0,1)\}^{y_{i3}^j}, \end{aligned} \tag{6}$$

and the unconditional likelihood $\mathscr{L}$ for a random sample of I consumers is

$$\mathscr{L} = \prod_{i=1}^{I} \int_{r_1}\int_{r_2}\int_{r_3} (\mathscr{L}_i \mid \mathbf{r}_i) f(r_{i1}, r_{i2}, r_{i3})\, dr_1\, dr_2\, dr_3, \tag{7}$$

where $f(r_{i1}, r_{i2}, r_{i3})$ is the density function of reservation prices across the population that we assumed to be distributed as multivariate normal $N(\bar{\mathbf{r}}, \boldsymbol{\Lambda})$. Equation (7) represents the likelihood function of a heterogeneous multinomial probit model where the price coefficient is set to one to fix scale indeterminacy and where the utility of the no-choice option is set to zero to fix origin indeterminacy.[7] Because price is ratio scaled, these constraints form the basis for obtaining dollarmetric reservation prices.

Before discussing model estimation, it is necessary to establish that the model parameters are identified. Rewriting Equation (2), we get

$$\begin{aligned}\Pr\{\mathbf{y}_i^j = (0,0,0)' \mid \mathbf{r}_i\} &= \Pr\{u_{i1j} \leq P_{1l_1} - r_{i1}, u_{i2j} \leq P_{2l_2} - r_{i2}, u_{i3j} \leq P_{3l_3} - r_{i3}\} \\ &= \Pr\left\{z_{i1j} \leq \frac{P_{1l_1} - r_{i1}}{\sqrt{\sigma_{11}}}, z_{i2j} \leq \frac{P_{2l_2} - r_{i2}}{\sqrt{\sigma_{22}}}, z_{i3j} \leq \frac{P_{3l_3} - r_{i3}}{\sqrt{\sigma_{33}}}\right\}, \end{aligned} \tag{8}$$

where σ_{mm}, $m = 1, 2, 3$, are the diagonal elements, respectively, of $\boldsymbol{\Sigma}$, and z_{i1j}, z_{i2j}, and z_{i3j} are standard normal variates.

Suppose prices P_{ml_m} are measured in dollars, say. Consider the probability of no purchase for any individual i given price scenario j. Suppose $(\mathbf{r}_i, \boldsymbol{\Sigma})$ and $(\mathbf{r}_i', \boldsymbol{\Sigma}')$ are two parameter solutions that satisfy $\Pr\{\mathbf{y}_i^j = (0,0,0) \mid \mathbf{r}_i\} = \Pr\{\mathbf{y}_i^j = (0,0,0) \mid \mathbf{r}_i'\}$. Then Equation (8) implies that $\mathbf{r}_i = \mathbf{r}_i'$ and $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}'$. Thus, it follows that $\mathbf{r}_i$ and $\boldsymbol{\Sigma}$ are identified. Consequently, $\bar{\mathbf{r}} = E(\mathbf{r}_i)$ and $\boldsymbol{\Lambda} = \mathrm{Var}(\mathbf{r}_i)$ are also identified. In addition, because prices are ratio-scaled, the estimated reservation prices are dollarmetric.

This measurement property provides important theoretical and managerial advantages. The firm can develop nonlinear pricing strategies and customized pricing plans for individual customers or customer groups. The social planner can compute consumer surpluses and hence unambiguously measure the social welfare implications of different product-line pricing policies.

[7] Consider the choice set $\bar{\Phi} = 0, 1, 2, 3$, where 0 denotes the no-purchase option. Define a latent variable $U_{imj} = r_{im} - \beta_i P_{mj} + \epsilon_{imj}$, $m = 0, 1, 2, 3$; $j = 1, \ldots, J_i$ where P_{mj} is the price of the mth choice alternative under price scenario j, β_i is the price coefficient and ϵ_{imj} is a random error. Then Equation (6) represents the likelihood function of a multinomial probit model where $u_{imj} = \epsilon_{imj} - \epsilon_{i0j}$, $(j = 1, 2, 3)$, $r_{i0} = 0$ (to fix origin indeterminacy), and $\beta_i = 1$ (to fix scale indeterminacy). Given these constraints, $U_{imj} = S_{imj}$, the latent surplus variable. Consequently, Equation (7) represents the likelihood function of a heterogeneous multinomial probit model with the aforementioned constraints. See the Appendix for a discussion of the equivalence betwen surplus maximization and utility maximization.

3.2. Model Estimation

The multinomial probit likelihood function in Equation (7) involves integration over high-dimensional multivariate normal distributions. Thus parameter estimation using maximum likelihood methods is difficult. We therefore use a hierarchical Bayesian framework to make inferences about the parameters based on their joint posterior distribution (see Allenby and Rossi 1998). In our case the unknowns are $(\{\{S_{imj}\}\}, \{\mathbf{r}_i\}, \bar{\mathbf{r}}, \mathbf{\Lambda}, \mathbf{\Sigma})$ where $S_{imj} = R_{imj} - P_{ml_m}$, is the surplus that consumer i derives from product/bundle m if it is priced at P_{ml_m}.

We use simulation-based methods to make many random draws from this posterior distribution. Inference is then based on the empirical distribution of this sample of draws. We use a substitution sampler which combines data augmentation (Albert and Chib 1993), the Gibbs sampler (Geman and Geman 1984), and the Metropolis-Hastings algorithm (Chib and Greenberg 1995). Substitution sampling replaces one complicated draw from the joint posterior with a sequence of relatively simple draws from easy-to-sample full conditional distributions of parameter blocks. This sequence of draws generates a Markov chain whose stationary distribution is the joint posterior density of all unknowns. Following the usual Markov chain Monte Carlo (MCMC) approach, the initial output from the chain is discarded because it reflects a transient or "burn-in" period in which the chain has not converged to the equilibrium distribution. A sample of draws is then obtained from the stationary distribution and used to make posterior inferences about model parameters and other quantities of interest.

For model estimation, we assume proper but non-informative priors. Details about the full conditional distributions we used in model estimation and the priors we specified are available on the *Marketing Science* Web site.

3.3. Model Performance

To assess model performance, we need a suitable measure of fit. For reasons described below, we use a methodology called "posterior predictive model checking" (Gelman et al. 1996). Traditional measures of goodness-of-fit such as the Bayesian information criterion (BIC) and consistent Akaike information criterion (CAIC) are not suitable for assessing and comparing hierarchical Bayes models because these measures are only asymptotically valid. In particular, such measures are inappropriate in situations where the number of model parameters varies with the sample size (Carlin and Louis 1996, p. 231). In addition, aggregate measures such as the log-likelihood statistics provide only limited information with respect to how well a model captures the data-generating process.

We carry out the posterior predictive model checks of internal validity using the Bayesian method described in Gelman et al. (1996). The method works as follows. First, fix the values of the explanatory variables to those that are observed in the sample. Then, use the posterior distribution of the model parameters to generate replicated datasets. These replicated datasets are predicted datasets based on random draws of the parameters from the posterior distribution. (This procedure is described in detail in the next paragraph.) If the replicated datasets differ systematically from the actual dataset on some specified dimension(s), conclude that the model is misspecified. An important advantage of this approach is that the posterior distributions can be generated during the MCMC sampling process itself. Consequently, no separate computations are necessary.

Formally, let y be the observed data, θ the vector of all unknown parameters, $p(y \mid \theta)$ the likelihood, and $p(\theta \mid y)$ the posterior distribution. For notational simplicity, these conditional probabilities do not explicitly refer to the design matrix because it is common across all replications. Suppose we obtain one thousand draws, $\theta_1, \theta_2, \ldots, \theta_{1000}$, from this posterior distribution as part of the MCMC sampling scheme. We now simulate 1,000 replications of the data, $y_1^g, y_2^g, \ldots, y_{1000}^g$, where y_i^g is drawn from the sampling distribution of y given θ_i. Thus, y^g has the posterior predictive distribution $p(y^g \mid y) = \int p(y^g \mid \theta) p(\theta \mid y) d\theta$. Comparisons of the replicated data with the actual data are carried out using measures known as *discrepancy variables* (Gelman et al. 1996). These discrepancy variables (described below) are chosen in light of the substantive questions the model seeks to answer. Thus a variety of discrepancy variables can be used to evaluate any particular model.

Recall that our primary goal is to predict a consumer's choice from a menu of options under a given price scenario. If a given model captures the data generation process well, then the frequency distribution of choices in the actual dataset should not be systematically different from the frequency distributions generated using the replicated datasets. As discussed earlier, under each price scenario, four mutually exclusive outcomes are possible. Thus, we seek to compare the actual frequency distribution of these four outcomes with the generated frequency distributions. A natural discrepancy variable is the difference between the number of times an alternative is chosen in the actual data and the number of times it is chosen using the generated data. In other words,

$$\mathrm{T}_i = \mathrm{C}_i - \mathrm{C}_i^g, \tag{9}$$

where T_i is the discrepancy variable for alternative i, $(i = 0, 1, 2, 3)$, C_i is the actual number of times that alternative i is chosen, and C_i^g is the generated frequency in a given replication i.

If the model overstates the number of times a specific alternative i is chosen, then T_i will be systematically greater than 0. If the model is "good" at capturing the data generating process, then the distribution of T_i should be symmetric around 0. In other words, on average, the model should predict the choice of alternative i the same number of times that it is present in the original data. Furthermore, the model should not systematically over- or underpredict this choice. Model performance can thus be gauged by plotting the distribution of T_i for each alternative i. A summary measure of model performance is then the p-value of $\mathrm{T}_i > 0$, which is calculated based on the empirical distribution of T_i. This p-value, which we denote by p_i, represents the proportion of replicated datasets for which $\mathrm{T}_i > 0$. Note that the ideal p-value for any alternative is 0.5.

A summary measure of the performance of each model can therefore be obtained using the following sum-of-squares deviation type performance statistic:

$$\mathrm{MP} = 1 - \frac{\sum_i (p_i - 0.5)^2}{\sum_i (0.5)^2}. \tag{10}$$

This measure varies between 0 ("bad" performance) and 1 ("good" performance). We will use the *MP* statistic to compare different models in our empirical studies.

4. Study Design and Data Collection

We carry out two related experimental studies to test our methodology. The first study has three main objectives. First, we test whether our methodology can capture the joint distribution of reservation prices. Second, we demonstrate how the model can be used to test theoretical assumptions. Third, we examine whether the model is useful for customer targeting (i.e., whether the model correctly identifies nonbuyers, buyers of the individual products, and buyers of the bundle).

The second study uses the same basic experimental design and product groups as the first study but focuses on three additional issues. First, we test the sensitivity of the method to the number of price scenarios each respondent evaluates. Specifically, each consumer in the first study evaluated nine price scenarios for each product group. In contrast, each consumer in the second study evaluated only three price scenarios for each of the corresponding product groups. Second, we compare our method with the simpler data collection strategy of asking consumers to directly state their reservation prices for individual products and bundles (the direct elicitation method). Third, we combine the choice-based and the direct elicitation methods to test whether self-stated prices provide any informational value beyond that provided by the choice-based method.

Subjects and Sample Size

The respondents in both studies were MBA students attending introductory marketing management classes at two large northeastern universities. The numbers of participants in the two studies were 88 and 78, respectively.

Stimuli

To assess model generalizability, we tested the model for a range of products, including both durables and nondurables. Specifically, in both studies, respondents were required to evaluate pricing scenarios for products in each of the following three product groups:

1. A videocamera (VC), a videocassette player/recorder (VP), and the bundle (VC-VP);
2. Six-month subscriptions to *Time* (TM) and *Newsweek* (NW) magazines, and a bundled subscription (TM-NW); and
3. A microwave oven (MO), a color television set (TV), and the bundle (MO-TV).

We chose these products for three reasons. First, the reservation prices vary considerably across products, ranging from low (for magazines) to high (for VP). Second, the interpurchase times also vary considerably across products, ranging from low for nondurables and high for durables. Third, our pre-test results showed that consumers were familiar with all these products.

Because a large number of brands of durables is typically available and each brand is available in a number of models, it was necessary to avoid ambiguity (across respondents) by specifying brand names and product attributes.[8] We chose well-known brands for each durable product; in addition, for each product we specified a set of product attributes taken from the *Consumer Reports Annual Buying Guide*. We specified brand names and product features as below:

- VC—Sony brand name, basic model, 10:1 power zoom lens, wireless remote feature, and date/time recording
- VP—Panasonic brand name, programming facility, auto head cleaner, front panel A/V inputs, 181 channel tuner, 1 year/8 event timer
- MO—Sharp brand name, 1 cubic foot capacity, 900 watts, 10 power levels, and 14.5 inch turntable
- TV—Sony brand name, color, 19-inch screen, Trinitron picture tube, stereo broadcast, and remote control

Choice of Price Scenarios

To obtain valid results from our experiments, it was necessary to choose credible price scenarios. We developed the scenarios using the following procedure. A pilot study was conducted using a convenience sample of 23 MBA students. Each subject was asked to state the maximum price that he/she would be willing to pay for each product. We then used the following three-step procedure. First, we eliminated outliers for each product and determined the following ranges of feasible prices—VC ($350 to $1,000), VP ($125 to $425), six-month subscriptions to TM and NW magazines ($15 to $60 for each), MO ($100 to $400), and TV ($175 to $500). In determining these ranges, we made sure that the market prices of these products fall within their respective ranges. Second, for each product we created three categories by dividing the appropriate range into three equal parts. For example, the three price categories for the microwave oven are as follows: $100 to $200 ("low"), $200 to $300 ("medium"), and $300 to $400 ("high"). Third, as explained in the next paragraph, we used the price categories for the individual products to construct the price categories for the bundle (i.e., low, medium, and high, respectively). Note that for plausibility it was necessary to ensure that, in any given scenario, the price of the bundle was no greater than the sum of the prices for the individual items. Hence, the definitions of "low," "medium," and "high" prices for the bundle were conditional on the prices chosen for the individual products.

Our procedure for specifying bundle prices in the price scenarios is best explained by example. Consider the magazine group and the scenario in which the prices for both the TM and NW subscriptions are low (say, $20 for TM and $25 for NW). Then the sum of the individual subscription prices is $45. Given this scenario, we used the following categorization scheme for specifying the price categories for the bundle. If the price of the bundle is 0 to 10% lower than $45 (i.e., in the range $40.50 to $45), the bundle price is high. If the bundle price is 10% to 20% lower than $45 (i.e., in the range $36 to $40.50), the bundle price is medium. Similarly, if the bundle price is 20% to 30% lower than $45 (i.e., in the range $31.50 to $36), the bundle price is low. Hence there were 27 cells corresponding to all combinations of prices (low, medium, and high) for the individual products and the bundle.

[8] In our experiment, the scenarios focus on price. As discussed in §7, our approach can be easily extended to construct scenarios where price, product features, and other product information vary.

Data Collection Procedure

We know a priori that each of the treatment variables (the prices of the individual products and the bundle) has a significant effect on choice behavior. Hence, a standard, completely randomized design is not statistically efficient (see Kirk 1982, p. 326). Furthermore, for practical reasons, it was necessary that each respondent evaluate only a limited number of price scenarios. We therefore used a hybrid approach that improves statistical efficiency by combining the full factorial methodology across all respondents and a Latin-square design for subgroups of respondents. The details of our data collection strategy in the first study are described below.

Consider any set of products and the bundle. Let "1," "2," and "3," respectively, refer to the low, medium, and high price treatments for the bundle. We used the following three-step procedure. First, randomly divide the sample a priori into three equal subgroups. Second, use a Latin-square design in which each respondent in each subgroup evaluates nine orthogonal price scenarios (see Table 1). Note that the design is fully randomized across all subjects (i.e., the price scenarios are mutually exclusive and exhaustive across subgroups). Third, to increase the validity of the experiment and to capture within-cell heterogeneity for any given subgroup, we used the randomization procedure described below.

Table 1 Latin-Square Design Used in Study 1

First Subgroup				
Price of		Product 1		
		Low	Medium	High
	Low	1	3[a]	2
Product 2	Medium	3	2	1
	High	2	1	3
Second Subgroup				
Price of		Product 1		
		Low	Medium	High
	Low	3	2	1
Product 2	Medium	2	1	3
	High	1	3	2
Third Subgroup				
Price of		Product 1		
		Low	Medium	High
	Low	2	1	3
Product 2	Medium	1	3	2
	High	3	2	1

[a]This treatment includes a medium price for product 1, a low price for product 2, and a high price for the bundle.

Consider the following treatment cell for respondents in the first subgroup—low prices for TM, NW, and the bundle (TM-NW). Recall that the relevant price ranges for the TM and NW subscriptions are \$15 to \$30. Consider two individuals, A and B, each of whom has been assigned randomly to the first subgroup. Then A's price scenario for the specified cell was determined as follows. Randomly choose prices for the TM and NW subscriptions from the appropriate ranges for the low price treatments. Suppose these prices are \$21 and \$24, respectively. Then the low price treatment for the bundle corresponds to the price range \$31.50 to \$36 (i.e., 20% to 30% less than the sum of the individual subscription prices). Now randomly choose a value from this range, say \$35. Then A's price scenario for this cell is \$21 (TM), \$24 (NW), and \$35 (the bundle). This procedure is repeated to construct B's price scenario. Clearly, these price scenarios will differ (i.e., the randomization strategy allows for within-cell heterogeneity in the price treatments).

When selecting the data collection procedure, it is important to recognize that repeated exposure of subjects to similar stimuli can threaten the internal validity of an experiment. Specifically, there is a testing effect if respondents become increasingly familiar with the task as the experiment proceeds. To minimize this problem, we required all respondents to perform a series of distraction tasks after evaluating each scenario. These distraction tasks ranged from completing a puzzle to finding the shortest distance on a road map for getting from one preassigned point to another.

The data collection procedure for the second study is similar to that described above. However, there are two important differences. First, in the second study each respondent only evaluated three scenarios for each product group. Consequently, we could not use the hybrid full factorial/Latin-square method. We therefore used a standard randomization approach across respondents—for any given product group, the set of price scenarios for each respondent is randomly

Table 2 Study 1—Estimated Reservation Price Distributions

Product Group	Product 1			Product 2			Bundle		
	$\bar{r}_1$	SD	SD/ $\bar{r}_1$	$\bar{r}_2$	SD	SD/ $\bar{r}_2$	$\bar{r}_3$	SD	SD/$\bar{r}_3$
VC-VP	561.81	89.00	0.158	231.21	62.89	0.272	694.19	100.19	0.144
TM-NW	22.43	11.77	0.525	23.73	12.01	0.506	33.63	15.89	0.472
MO-TV	157.69	67.34	0.427	264.40	74.73	0.283	418.65	83.12	0.199

Note. $\bar{r}_m$ = Mean reservation price for product m; SD = posterior standard deviation of reservation price, $\sqrt{(\lambda_{mm})}$.

chosen from the full set of 27 price scenarios. Note that, as in the first study, we used a randomization procedure to allow for within-cell heterogeneity to improve statistical efficiency. Second, to compare our model to the direct elicitation method, we required respondents to state their reservation prices for the appropriate products and bundles. The subjects stated their reservation prices for each of the three products/bundles prior to their participation in the choice task.

5. Results and Analysis

We estimated all models using code written in the C-programming language. As discussed in §3.2, we used the MCMC methodology to make repeated draws from the series of full conditionals to arrive at the joint posterior density of the unknown parameters. The substitution sampler was run for 10,000 iterations for each model. We examined the sequence of output draws to ensure that convergence had taken place and used the last 5,000 draws for making inferences.

5.1. Study 1

We first examine the descriptive results and then discuss the managerial implications for bundle pricing.

5.1.1. Heterogeneity. Table 2 reports the summary statistics for the distributions of reservation prices for the individual products and bundle for each product group. The main results are fourfold. First, for all product groups there are considerable differences across consumers' reservation prices. Second, the reservation prices for the nondurable products (TM and NW) are considerably more heterogeneous than those for the durable products used in the study. Specifically, the coefficients of variation[9] for TM and NW, respectively, are 0.525 and 0.506—in contrast, except for MO, the highest coefficient of variation for the durables in the study is only 0.283. Third, regardless of whether the products are durables or nondurables, bundling reduces the degree of demand heterogeneity.[10] For example, the coefficient of variation for the MO-TV bundle (0.199) is less than the corresponding value for TV (0.283). Fourth, for all products and bundles, the within-subject variations are significant—specifically, the lowest ratio of σ_{mm} to its standard deviation is 3.25 (Table 3). However, as a comparison of Tables 2 and 3 shows, between-subject heterogeneity in reservation prices accounts for most of the variability. Recall that within-subject variation captures idiosyncratic randomness in choice due to consumers' error in judgment, whereas between-subject variation measures the dispersion in the true reservation prices across individuals.

Table 3 Study 1—Within-Subject Reservation Price Variances

Product Group	σ_{11}		σ_{22}		σ_{33}	
	Mean	SD[1]	Mean	SD	Mean	SD
VC-VP	0.88	0.18	0.47	0.11	1.08	0.22
TM-NW	1.88	0.34	1.47	0.27	2.08	0.57
MO-TV	1.77	0.29	1.39	0.31	2.15	0.66

Note. SD = posterior standard deviation.

5.1.2. The Joint Distribution of Reservation Prices. We begin by testing the assumption that

[9] We use the coefficient of variation as our measure of heterogeneity of reservation prices to compare our results across the three experiments. Note that this measure is scale-free.

[10] This result follows trivially when reservation prices are additive. However, it does not necessarily hold when reservation prices are subadditive or superadditive.

Table 4 Model Comparisons for Testing Individual-Level Additivity of Reservation Prices

	Hit Rate Comparison				MP Statistic Comparison	
	Full Sample		Holdout Sample		Full Sample	
Product Pair	Unrestricted	Restricted	Unrestricted	Restricted	Unrestricted	Restricted
VC-VP	78	69	61	44	0.96	0.85
TM-NW	93	83	85	67	0.85	0.73
MO-TV	81	71	68	49	0.87	0.76

reservation prices are additive (e.g., Schmalensee 1984, p. S211). Individual-level additivity implies that additivity holds at the population level (i.e., the mean reservation price for the bundle is the exact sum of the mean reservation prices for the individual products). However, it is important to note that additivity at the population level—a crucial assumption driving the theoretical results on bundling—does not imply individual-level additivity.[11]

To test for individual-level additivity, we estimated a restricted version of our model in which, for any given individual, the reservation price for the bundle on a given choice occasion is the exact sum of the appropriate estimated reservation prices for the products in the bundle. We then compared the unrestricted and restricted models using two criteria: (1) the hit rates for the full sample and the holdout sample and (2) the MP discrepancy statistics (see §3.3).

As shown in Table 4, for all three product groups, the hit rates for the unrestricted model (full and holdout samples) are considerably higher than those for the individual-level additivity (restricted) model. The MP statistics also show that the restricted model performs poorly in capturing the data-generating process. Thus, the assumption of individual-level additivity is not satisfied in any of the three experiments.

We now examine whether the additivity assumption holds at the population (i.e., mean) level (See Table 2). The results show that population-level additivity is supported ($p < 0.05$) for only one of the three product groups (MO-TV). For the other two product groups (VC-VP and TM-NW), the mean reservation prices are strongly subadditive. In particular, for the VC-VP group, the mean reservation price for the bundle is 13% below the sum of the mean reservation prices of VC and VP. For the TM-NW group the mean reservation price for the bundle is 40% below the sum of the mean reservation prices of TM and NW.

In all cases the assumption of statistical independence of reservation prices is strongly rejected. Specifically, the highest correlation is for the VC-VP pair (0.89; standard error = 0.03) followed by the TM-NW pair (0.72; standard error = 0.04) and the MO-TV pair (0.51; standard error = 0.07).

Interestingly, even if additivity holds at the population level (as in the case of the MO-TV pair), the assumption of statistical independence may not. Thus, models that assume statistical independence are likely to be misspecified. Importantly, using individual-level estimates of the reservation prices of individual products to construct the reservation prices of bundles is likely to lead to erroneous conclusions.

5.1.3. Model Performance. The previous analysis implicitly assumes that the correct model must incorporate heterogeneity. However, before using the results of our model, it is important to test whether this model performs better than a model without heterogeneity. We therefore contrast the performance of our model with two models that assume homogeneity. The first is a nested version of our model where $\mathbf{r}_i = \bar{\mathbf{r}}$ for all i and $\boldsymbol{\Lambda} = \mathbf{0}$. The second is a model based on the approach used in GGW.

We use two tests to evaluate model performance. First, we use the *MP* discrepancy statistic (see §3.3). Second, we compare the choice prediction abilities of the three models using hit rates. In this analysis, we begin by comparing the hit rates for all three models using the full-sample data. Then, to test for robustness, we re-estimate the hit rates for all three models

[11] We thank an anonymous referee for drawing our attention to this point.

using a holdout dataset that contains one randomly chosen observation for each respondent (recall that each respondent evaluated only nine price scenarios). Note that each model classifies an observation into one of four mutually exclusive categories: no purchase, choice of product 1, choice of product 2, and choice of the bundle.

Comparison with the Nested Model

The *MP* statistics show that the fit improves considerably when the model includes heterogeneity. For the VC-VP product bundle, *MP* increases from 0.68 in the nonheterogeneous model to 0.96 in the heterogeneous model. For the TM-NW and MO-TV product bundles, respectively, the corresponding increases are from 0.55 to 0.85, and from 0.69 to 0.87. Recall that the closer the *MP* statistic is to 1.00, the better the model fit.

Table 5 contains the hit rates using both the full and the holdout samples. Note that "Het. RP" refers to heterogeneous reservation prices and "Non-Het. RP" refers to nonheterogeneous reservation prices. We first discuss the full-sample results.

The full-sample results show that, for all the three product groups, the nonheterogeneous model performs better than chance in predicting purchase of the individual products and non-purchase. However, regardless of the product group, this model consistently fails to perform significantly better than chance in classifying buyers of bundles. The hit rate for the TM-NW bundle in the nonheterogeneous model is especially poor (3%). In contrast, the heterogeneous model always performs better than chance regardless of the product group or the alternative that is chosen. Interestingly, and in sharp contrast to the nonheterogeneous model, the heterogeneous model performs especially well for nondurables. Specifically, the hit rates for magazines range from 97% for the no-purchase alternative to 55% for the choice of the bundle.

The results for the holdout sample reinforce the previous findings. For example, in the videocamera study, the hit rate for the nonheterogenous model (19%) is not significantly different from chance. Furthermore, the hit rate for nonbuyers is zero. In contrast, the corresponding rates for the heterogeneous model (61% and 48%, respectively) are significantly better than chance. The results for the other two product groups are similar, with the heterogeneous model performing particularly well for nondurables. Specifically, the hit rates for nonbuyers and buyers of the TM-NW bundle, respectively, are 85% and 83%.

Table 5 Study 1—Hit Rates (%)

	Full Sample			Holdout Sample		
	Het. RP	Non-Het. RP	Het. GGW[a]	Het. RP	Non-Het. RP	GGW[a]
No buy	82	60	40	48	0	100
VC	73	73	61	83	33	0
VP	80	61	70	33	27	0
VC-VP	63	32	32	94	9	0
Overall	78	57	50	61	19	24
No buy	97	79	48	85	9	100
TM	90	41	73	91	30	0
NW	89	45	66	79	79	0
TM-NW	55	3	37	83	17	0
Overall	93	64	56	85	31	33
No buy	88	70	41	59	3	100
MO	77	39	54	67	0	0
TV	78	46	61	37	5	0
MO-TV	60	32	55	97	90	0
Overall	81	57	53	68	34	25

[a] Results using the GGW method.

Comparison with the GGW Method

The GGW method provides reasonably good predictive results, as measured by the hit rates for the estimation sample (see Table 5, Column 4). However, these results are inferior to those from the heterogeneous model. To test whether the GGW results are valid, we examined the canonical roots from the GGW method. For all datasets, the second canonical roots were highly significant ($p < 0.0001$). As GGW note (p. S130), "Of course, if the second [or any other] set of linear compounds had turned out to yield large and significant correlations, doubt is cast on the appropriateness of the model." Thus, the GGW method is inappropriate for our data.

This lack of robustness was confirmed by the holdout analysis. For all data sets, the GGW model fails to perform better than chance (see Table 5, Column 7). Furthermore, regardless of whether the products are durables or nondurables, the model classified all observations in the holdout sample into the no-purchase category.

5.1.4. Conclusion. The results show that for both durables and nondurables, reservation prices are heterogeneous. Our model predicts well, regardless of whether the subjects are nonbuyers, buyers of individual products, or buyers of the bundle. Hence, our model can be used for targeting consumers. For all three product groups, the heterogeneity across consumers for the product bundle is less than the corresponding heterogeneities for the individual products. Finally, empirical methods that assume individual-level additivity and independence across reservation prices are likely to be misspecified.

5.2. Study 2

5.2.1. Effect of the Number of Profiles per Respondent. Table 6 (Columns 5–7) contains the summary statistics for the distributions of the estimated reservation prices for the different products and product bundles. Comparing Tables 2 and 6 we see that the means and standard deviations of the estimated reservation prices do not appear to be affected by the number of profiles. For example, the mean estimated reservation prices for the videocamera are \$561.81 (Study 1) and \$572.75 (Study 2) and the corresponding standard deviations are \$89.00 and \$81.33. Formal statistical tests (not reported here) show that the number of profiles per respondent does not have a significant effect on the estimated distributions of reservation prices. The *MP* discrepancy statistics (see Equation (10)) show that there is a marginal improvement in fit when the number of profiles is increased. For the VC-VP product group, the *MP* statistic increases from 0.92 to 0.96 when the number of profiles per subject increases from three to nine. The corresponding values for the other two product groups are 0.82 and 0.85 (TM-NW) and 0.81 and 0.87 (MO-TV), respectively. Thus, the methodology appears to be robust even when consumers evaluate a limited number of profiles.

Table 6 Study 2—Summary Statistics for Self-Stated and Estimated Reservation Prices

Product/ Bundle	Self-Stated			Estimated		
	RP[a]	SD[b]	SD/RP	RP[c]	SD[d]	SD/RP
VC	554.17	257.25	0.46	572.75	81.33	0.14
VP	246.19	93.08	0.38	235.69	53.02	0.23
VC-VP	722.60	279.89	0.39	701.40	94.45	0.14
TM	22.81	9.28	0.41	21.90	10.97	0.50
NW	22.39	9.06	0.41	23.12	11.66	0.50
TM-NW	36.77	15.75	0.43	31.29	14.45	0.46
MO	187.65	62.94	0.34	160.47	63.44	0.38
TV	303.71	117.82	0.39	260.34	64.58	0.25
MO-TV	438.41	134.36	0.31	424.65	77.21	0.18

[a] Mean of self-stated reservation prices.
[b] Standard deviation of self-stated reservation prices.
[c] Mean of estimated reservation prices.
[d] Posterior standard deviation of estimated reservation prices, $\sqrt{\lambda_{mm}}$.

5.2.2. The Choice-Based and Direct Elicitation Methods. We begin by comparing the summary statistics for the choice and direct elicitation methods (see Table 6). The results show that although the mean reservation prices are very similar for both models, the standard deviations are very different. For the durable product groups, the estimated standard deviations of the reservation prices for the direct elicitation method are considerably higher than for the choice-based method. For example, in the VC-VP group, the estimated standard deviations of reservation prices for the bundle are \$279.89 (direct elicitation method) and \$94.45 (choice model). In contrast, for the nondurable product group (magazines), the estimated standard deviations of reservation prices for the direct elicitation method are somewhat lower than for the choice-based method.

These results show that, at least for durables, the direct elicitation method implies that the demand for individual products and bundles is more heterogeneous than for the choice-based method. That is, the estimated demand curves for both methods are very different, even though the mean reservation prices are not. For example, if the price of a bundle of durables is higher than the mean reservation price for that bundle, the direct elicitation method implies that demand is less elastic than is implied by the choice-based method (see Schmalansee 1984, p. 214). Hence, the managerial implications for product-line pricing can be significant.

We now compare the hit rates for both methods (see Table 7). The results show that, for all the three product groups, the choice-based method outperforms the direct elicitation method by a margin ranging from 8% (VC buyers) to 43% (NW buyers). Importantly,

Table 7 Study 2 - Hit Rates (%)

	Choice-Based RP	Self-Stated RP
No buy	83	51
VC	58	50
VP	78	60
VC-VP	54	43
Aggregate	72	51
Aggregate	96	86
TM	96	53
NW	89	46
TM-NW	78	33
Aggregate	93	68
No buy	94	71
MO	90	72
TV	82	74
MO-TV	59	20
Aggregate	84	61

from a targeting and resource allocation viewpoint, the direct elicitation method does very poorly in identifying buyers of bundles. Specifically, for two of the three product groups (TM-NW and MO-TV), the hit rates for the bundles (33% and 20%, respectively) are no better than chance.

These results have significant implications. Using self-stated reservation prices leads to highly biased demand estimates, particularly for durables. In particular, this method performs extremely poorly in predicting demand for the bundle. Consequently, optimization algorithms that use these values as inputs are likely to lead to suboptimal decisions.

5.2.3. The Augmented Model. We now examine whether an augmented, choice-based model that includes self-stated reservation price information performs better than the heterogeneous reservation price model in Equation (1). Specifically, we reparameterize the true reservation prices r_{im} in Equation (1) as functions of the self-stated reservation prices ss_{im}. That is, $r_{im} = a_{im} + b_m \times ss_{im}$, where a_{im} is an individual-specific intercept and b_m is an adjustment factor that reflects the differential informational value of self-stated reservation prices. Thus, an adjustment factor of unity corresponds to the special case where consumers can state their reservation prices with perfect accuracy and an adjustment factor of zero corresponds to the case where self-stated reservation prices have no informational content. With this reparameterization, the augmented reservation price model is

$$R_{imj} = a_{im} + b_m \times ss_{im} + u_{imj},\ m = (1, 2, 3);$$
$$i = 1, \ldots, I;\ j = 1, \ldots, J, \quad (11)$$

where the vector $(a_{i1}, a_{i2}, a_{i3})'$ is distributed as multivariate normal, $N(\bar{\mathbf{a}}, \mathbf{\Lambda})$. Note that if $b_m = 0; m = 1, 2, 3$, then the augmented reservation price model in Equation (11) reduces to the basic reservation price model in Equation (1).

Table 8 contains the estimated adjustment factors, b_m, and their posterior standard deviations. Now the ratios b_m/SD denote approximate "t-statistics." Using a one-tailed "t-test," we see that except for the MO-TV bundle, all the adjustment factors are statistically significant ($p < 0.05$). Thus, in general, the self-stated reservation prices do have some informational content in the augmented model. However, this effect is only strong for the nondurable product group (TM-NW) where the adjustment factors range from 0.642 for the bundle to 0.770 for TM. In contrast, for the durable product groups, the adjustment factors are small, ranging from 0.064 to 0.193.

These conclusions are reinforced by comparing the within-sample predictive abilities of the two models. For two of the three product pairs (VC-VP and TM-NW) the hit rates increase when the augmented model is used. However, the improvement

Table 8 Study 2—Estimated Adjustment Factors, b_m, in Augmented Model

Product/Bundle	b_m	SD	$\frac{b_m}{SD}$
VC	0.064	0.036	1.78
VP	0.136	0.058	2.34
VC-VP	0.085	0.031	2.74
TM	0.770	0.187	4.12
NW	0.755	0.173	4.36
TM-NW	0.642	0.201	3.19
MO	0.193	0.106	1.82
TV	0.013	0.057	2.28
MO-TV	0.028	0.055	0.51

Note. SD denotes posterior standard deviation.

is marginal. Specifically, for the VC-VP pair, the hit rate increases from 72% to 77% when the augmented model is used. The corresponding increase for the magazine study is smaller (from 93% to 95%). However, for the MO-TV group, there is no improvement at all (the hit rates for both models are 84%).

Three main results emerge from Study 2. First, our model captures heterogeneity reasonably well even when respondents evaluate a limited number of profiles. Second, for all the three product groups examined, our method outperforms the direct elicitation method in predicting choice. Third, the informational content in self-stated reservation prices is marginal, at best. Consequently, in general, it may not be efficient to use the augmented model unless the cost of collecting self-stated data is low.

6. Managerial Implications

In this section, we show how the firm can use the results from our model to choose the optimal product-line pricing policy for the individual products and the bundle. We contrast this analysis with an optimization analysis based on self-stated reservation prices.

6.1. Optimization Results Using Our Model

Consider any product-line pricing policy (P_1, P_2, P_3). Without loss of generality, consider the proportions of consumers who buy the individual products and the bundle. Let z_1, z_2, and z_3, respectively, denote the proportions of consumers who buy products 1, 2, and the bundle. Assume that the marginal cost of production for each product is constant (this assumption can be relaxed). Let C_1, C_2, and C_3 denote the marginal production costs for products 1, 2, and the bundle, respectively. Then, ignoring fixed costs, the firm's profit, π, under this price policy is given by

$$\pi = (P_1 - C_1) * z_1 + (P_2 - C_2) * z_2 + (P_3 - C_3) * z_3. \quad (12)$$

The optimal prices (P_1^*, P_2^*, P_3^*) can be calculated by differentiating this profit function with respect to P_1, P_2, and P_3, setting the derivatives equal to zero, and solving for the three prices. However, given that z_1, z_2, and z_3 are stochastic (they are estimated using Equations (3)–(5) inclusive), the optimal solution is difficult to obtain analytically (see also Schmalansee 1984, p. S225). Thus, it becomes necessary to use an algorithm-based approach to compute profit.

We now develop and implement an algorithm for choosing optimal product-line pricing policy given the estimated reservation prices (and, hence, the z_is). Our approach makes two assumptions, both of which may be relaxed if necessary.

First, because the optimization involves parameter estimates (i.e., the true parameters that describe the reservation prices are unknown), it is necessary to specify the firm's risk attitude. We assume that the firm is risk-neutral (i.e., the firm chooses policy to maximize expected profits). This objective function can be changed if the firm is risk-averse and we know its risk-return trade-off.

Second, as mentioned earlier, there are no economies of scale or scope in production. Because specific cost data are not available, we used approximate product-category level information derived from *Advertising Age* (June 29, 1998, p. 22). Specifically, we constructed cost ratios using the reported values for the advertising-sales and advertising-margin ratios for the appropriate product categories. The following cost ratios were obtained: 51% (periodicals), 71% (household audio/video), and household appliances (70%). We used the appropriate values for each product analyzed. For example, the cost ratios for VC and VP are 71%. To determine marginal costs, it is necessary to combine the cost ratios with an appropriate set of prices for the individual products and bundles. In our analysis, we used the mean market prices. To illustrate, the mean price for the TM subscription package was $37 and the cost ratio for periodicals was 0.51. Hence the estimated marginal cost for TM is $18.87.

We now detail the steps required to compute profits:

1. Choose a product-line pricing scenario $(P_1, P_2, P_3)_l$ from $P_1 \times P_2 \times P_3$.
2. Use the parameter estimates to draw the reservation prices for all consumers in the sample.
3. Compute the proportions of consumers who will choose each of the available alternatives (no purchase or purchase of products 1, 2, or the bundle) given the specified product-line pricing policy.

4. Use Equation (12) to compute profits, given the cost structure.

5. Repeat Steps 2 through 4 one hundred times to estimate the distribution of profits for this pricing policy.

6. Compute the mean profit and the associated standard deviation under this pricing policy.

7. Return to Step 1 and evaluate all other candidate product-line pricing policies.

8. Choose the product-line pricing policy that maximizes expected profit.

The results show that a mixed bundling strategy is optimal for all the three product groups. However, there are important differences across the three product groups. For the VC-VP product pair, the optimal product-line pricing policy is as follows: \$520 (VC), \$256 (VP), and \$670 (bundle). This policy provides a discount of 13.66% to buyers of the bundle and produces an expected profit of \$66.75 with a standard deviation of \$0.65. For the TM-NW group, the optimal pricing strategy (\$38 for TM, \$40 for NW, and \$70 for TM-NW) provides a discount of 10.26% to those who buy the bundle and produces an expected profit of \$5.55 with a standard deviation of \$0.19. For the MO-TV group, the optimal pricing strategy (\$235 for the MO, \$314 for the TV and \$510 for MO-TV) provides a 7.10% discount to those who purchase the bundle and yields an expected profit of \$31.48 with a standard deviation of \$0.48.

Define a high-price strategy as charging a price higher than the mean reservation price for a given product or bundle. The results show that it is not always optimal to use a high-price strategy to extract consumer surplus. For the magazine study, the mean reservation prices (\$22.31 for TM, \$23.60 for NW, and \$33.11 for TM-NW) are considerably lower than the corresponding optimal prices (\$38 for TM, \$40 for NW, and \$70 for TM-NW bundle). Thus, a high-price strategy is optimal. This result is obtained because the coefficients of variation for the products/bundle in this group are very high, ranging from 0.472 to 0.525 (see Table 2). Consequently, the demands for each magazine and the bundle are highly inelastic for prices larger than the corresponding mean reservation prices. It follows that charging high prices for the entire product line (i.e., the individual products and the bundle) is an efficient method of extracting consumer surplus.

For the MO-TV group, the demands for the individual products and the bundle are also relatively heterogeneous, with coefficients of variation ranging from 0.199 to 0.427. Hence, as in the magazine study, a uniform high price strategy is optimal. Specifically, the mean reservation prices (\$158.30 for MO, \$263.50 for TV, and \$419.98 for the bundle) are considerably lower than the corresponding optimal prices (\$235 for the MO, \$314 for TV, and \$510 for the bundle).

In contrast, for the VC-VP pair, demand is less heterogeneous. Specifically, the coefficients of variation are as follows: 0.158 (VC), 0.272 (VP), and 0.144 (the bundle). Consequently, although a high-price strategy is optimal for VP (which has the highest coefficient of variation), this strategy is suboptimal for VC and the bundle. Specifically, the optimal prices for VC and the bundle, respectively, are 7.84% and 3.71% below the corresponding mean reservation prices.

6.1.1. Optimization Results Using Self-Stated Reference Prices. The optimization analysis using self-stated reservation prices can be conducted in two ways. First, one can assume continuous heterogeneity (as in our model) and directly use the individual-level, self-stated reservation prices to measure demand. Alternatively, one can assume that heterogeneity is discrete (segment-level) and use the self-stated reservation prices to form segments. We performed both types of analysis.

In the latter case, we used the approach described in Hanson and Martin (1990). We created the segments as follows. In the first step, we formed discrete clusters (segments) using the self-stated reservation prices. We used the Ward clustering method and chose the number of clusters for each dataset using the scree test and the R^2 statistics. The optimal numbers of clusters were as follows: three (VC-VP), five (TM-NW), and three (MO-TV). The R^2 values were as follows: 0.87 (VC-VP), 0.84 (TM-NW), and 0.84 (MO-TV). In the second step, we ran an optimization algorithm assuming that the reservation prices of all consumers in any given segment are located at the centroid for that group (i.e., we assumed that all within-segment variabilities are zero).

Table 9 Profitability Comparisons (in $) Across Methods

Product Group	Optimal Prices			Profit	SD[a]
	Product 1	Product 2	Bundle		
VC-VP					
Our method	520	256	670	66.75	0.65
Self-stated (segment-level)	570	267	785	62.80	
Self-stated (individual-level)	761	250	924	58.71	
TM-NW					
Our method	38	40	70	5.55	0.19
Self-stated (segment-level)	40	26	55	3.86	
Self-stated (individual-level)	40	38	62	2.88	
MO-TV					
Our method	235	314	510	31.48	0.48
Self-stated (segment-level)	300	291	500	38.66	
Self-stated (individual-level)	273	334	576	33.70	

[a]SD denotes standard deviation of profit. Note that profits are deterministic for methods that use self-stated reservation prices.

Table 9 shows the profits for the optimization carried out using self-stated reservation prices. Note that, in contrast to our method, which yields a stochastic estimate of profits, using self-stated reservation prices leads to deterministic profit values. Also, recall our assumption that the firm is risk-neutral. Hence, all comparisons with our methodology are based on the expected profits from the relevant optimal policies.

The results show that if heterogeneity is continuous, using individual-level, self-stated reservation prices leads to very different pricing strategies and profits from our model. For the VC-VP study, the self-stated method leads to significantly higher prices than our model (e.g., the bundle price is 37.91% higher). This result is not surprising because the self-stated method drastically overstates the heterogeneity in reservation prices (e.g., the standard deviation of reservation prices for the VC-VP bundle is almost three times larger than the corresponding value for our model). Hence, it is optimal for the firm to uniformly charge high prices across the product line to extract consumer surplus. The results for the self-stated method also suggest that profits will be 12.04% lower than those from our model.

For the MO-TV study, the standard deviations of the reservation prices for MO are almost identical across methods ($62.94 and $63.44, respectively, for the self-stated method and our model). However, the self-stated method significantly overstates the heterogeneity in reservation prices for TV and the MO-TV bundle (in each case by a factor of approximately two). Consequently, as in the VC-VP study, the self-stated method leads to higher prices for the individual products (MO and TV) and the bundle. However, in contrast to the VC-VP study, the self-stated method implies that profits will be higher (by 9.68%) than profits in our model.

The TM-NW results are interesting. For TM and NW, the self-stated method provides lower estimates of heterogeneity for the individual products than our method. However, the opposite is true for the bundle. Consequently, the pricing implications are more complex: The self-stated method leads to higher prices for TM but lower prices for NW and the bundle. In addition, the self-stated method implies that profits will be 48.11% lower than profits using our model.

We now compare the results for the two heterogeneity scenarios (continuous and discrete) using self-stated reservation prices. The results show that, for all product groups, the discrete (i.e., segment-level) heterogeneity specification leads to somewhat higher profitability than the continuous one (i.e., individual-level). The improvements in profitability are as follows: 34.06% (TM-NW), 14.72% (MO-TV), and 6.97% (VC-VP).

Finally, we compare the results from our model to those from the discrete heterogeneity specification. No clear pattern emerges for either the price or profit comparisons. This result is not surprising. Recall that in our experiments self-stated reservation prices were found to contain significant measurement error. Furthermore, clustering algorithms perform very poorly when the data contain measurement error. See the extensive simulation results in Milligan (1980).

6.2. Discussion

The empirical results show that it is difficult to make broad general statements regarding optimal product-line pricing policy. Although mixed bundling was optimal in all cases, there were significant differences across product categories. The main finding is that a uniform high-price policy is likely to be optimal when the reservation prices for the individual products and the bundle are very heterogeneous. However, when the demand for a product or the bundle is relatively homogeneous, this policy is likely to be suboptimal. In such cases, it may be necessary to charge lower prices for some products or bundles in order to capture consumer surplus.

Using self-stated reservation prices as an approximation for the true (unobservable) reservation prices leads to significantly lower profitability than the optimal strategy. However, the effect on product-line pricing policy cannot be specified a priori. This ambiguity is not surprising. Recall that our internal validity checks for all the product categories show that self-stated reservation prices are biased. Consequently, regardless of whether one assumes continuous or discrete heterogeneity, the market or segment-level demand curves derived from self-stated reservation prices are also biased.

7. Model Extensions

The proposed model is highly flexible and can be extended to analyze many different scenarios. We consider four important cases: (1) the effect of product design on reservation prices, (2) the formation of individual-level reservation prices, (3) the impact of competitive products and/or product information, and (4) more complex bundling situations.

Suppose the manager seeks to determine how product design affects reservation prices. Similarly, suppose the researcher has a theory of antecedents (e.g., previous purchase history, internal or external reference prices) that affect the true reservation prices. Then, following our approach in Equation (11), one can rewrite Equation (1) as a function of the product design features and the relevant antecedent factors. Let $\mathbf{x_m}$ be a $(p_m \times 1)$ vector of product design features for alternative m. Let $\mathbf{z_i}$ be a $(q \times 1)$ vector of antecedent variables for consumer i. Then for product alternatives $m = 1, 2$,

$$R_{im} = \bar{r}_{0m} + \boldsymbol{\alpha}'_m \mathbf{z}_i + \boldsymbol{\beta}'_{im} \mathbf{x}_m + u_{im}, \quad um = (1, 2);\ i = 1, \ldots, I, \tag{13}$$

where $\bar{r}_{0m}$ is an alternative-specific intercept, $\boldsymbol{\alpha}_m$ is a $(q \times 1)$ vector that measures the effect of antecedent factors, $\boldsymbol{\beta}_{im}$ is an individual-specific $(p_m \times 1)$ vector that measures the effect of product design features, and u_{im} is an error term. Following Venkatesh and Kamakura (2003) the bundle $(m = 3)$ reservation price can be specified as

$$R_{i3} = (1 + \gamma_i) \times (r_{i1} + r_{i2}) + u_{i3}, \quad i = 1, \ldots, I, \tag{14}$$

where $r_{im} = \bar{r}_{0m} + \boldsymbol{\alpha}'_m \mathbf{z}_i + \boldsymbol{\beta}'_{im} \mathbf{x}_m$ and γ_i is an individual-specific coefficient that measures the departure of the bundle reservation price from additivity. Thus, $\gamma_i < 0$ $(\gamma_i > 0)$ implies subadditivity (superadditivity). To model unobserved consumer heterogeneity, we assume that the joint vector $(\boldsymbol{\beta}'_{i1}, \boldsymbol{\beta}'_{i2}, \gamma_i)'$ is distributed across the population of consumers as multivariate normal $N((\bar{\boldsymbol{\beta}}'_1, \bar{\boldsymbol{\beta}}'_2, \bar{\gamma})', \boldsymbol{\Lambda})$, where $(\bar{\boldsymbol{\beta}}'_1, \bar{\boldsymbol{\beta}}'_2, \bar{\gamma})'$ is the mean vector and $\boldsymbol{\Lambda}$ is the covariance matrix. This extended model can be estimated using a Bayesian approach similar to that used for estimating the augmented model (see §5.2.3).[12]

Suppose the manager seeks to measure the impact of the firm's product-line policy on competitors' performance. Alternatively, suppose the manager seeks to measure the impact of competitors' product-line pricing policy on the firm's performance. Then the

[12] We were unable to estimate the extended model in this study because the experimental design did not manipulate product features or control for antecedent variables.

only modification necessary is to include the relevant competitive products in the experiment. Note that because of the inclusion of the no-purchase option in the choice set, omitting competitive products from the experiment does not lead to biased estimates of the joint distribution of the reservation prices of those products or bundles that are included in the experiment (see §3).

Suppose the manager seeks to determine the effect of product information on consumers' reservation prices and choices. Then the scenarios in the experiment should consist of the relevant price-information combinations. Note that the informational set can be very general. For example, it can include information about competitive alternatives not considered in the experiment or for products that do not yet exist.

Finally, consider more complex bundling situations. In our empirical application we tested a simple model (no purchase, two products, and one bundle). Theoretically, as discussed in §3, the model can be used to estimate models containing any number of alternatives (products and/or bundles). Suppose the firm's product line contains three products. Then a full experiment will contain seven alternatives, including four bundles. Recall that the model appears to be robust even when the number of profiles per subject is small. Consequently, we expect the model to perform well in capturing the joint distribution of reservation prices. Suppose the firm's product line contains more than three products (say four). Then a full experiment will contain fifteen alternatives. In such cases, it may be necessary to use specialized experimental designs. See, for example, the methods discussed in Louviere and Woodworth (1983) in the context of conjoint studies. As the number of products in the firm's product line increases beyond four, the data required to obtain robust parameter estimates will increase rapidly. Hence the method may become impractical. However, as casual empiricism shows, bundles that are sold in the market typically contain less than four products. Hence our method should be useful for most practical applications.

8. Discussion and Summary

This paper develops and tests a model for estimating the joint distribution of reservation prices for products and bundles. In addition, we develop and test an algorithm for using the model estimates to choose the optimal product-line pricing policy.

Our approach provides several methodological and managerial advantages. The model is derived from utility theory and captures continuous heterogeneity. In addition, it measures both within- and among-subject variability without imposing any structure on the joint distribution of reservation prices for the individual products and bundles. Hence the model allows for general forms of substitutability and complementarity and can be used to test theoretical assumptions. The model provides dollarmetric reservation prices and allows the user to obtain individual-level estimates. Hence the manager can use the model estimates to target customers, choose optimal product-line pricing policy, and develop customized (including nonlinear) pricing policies. In addition, planners can use the model to determine the social policy implications of bundling pricing policies in different industries.

We conduct two experimental studies to test model robustness. To assess generalizability, each study examines three product groups containing both durable and nondurable products. The results show that, regardless of product group, our model captures heterogeneity and performs well in terms of fit and predictive accuracy. In contrast, models that do not allow for heterogeneity perform poorly, especially in predicting bundle choice.

Our experiments support the theoretical assumption that reservation prices for the bundle are less heterogeneous than those for the individual products in the bundle. However, they do not support the assumptions that individual-level reservation prices are additive and statistically independent.

Our methodology appears to be robust even when respondents evaluate only a limited number of profiles. Regardless of the type of product (durable or nondurable), our model is considerably superior to the direct elicitation method, in which respondents directly state their reservation prices for individual products and bundles. In particular, the direct elicitation method performs very poorly in predicting bundle choice. In general, self-stated reservation prices do not have any informational content beyond that contained in the basic heterogeneity model. The direct elicitation method appears to understate (overstate) the variation in reservation prices across consumers for low-priced (high-priced) products and bundles. Hence, this method yields biased estimates of the appropriate demand curves and leads to suboptimal product-line pricing policy.

The optimization results show that the self-stated approach for measuring reservation prices leads to a suboptimal pricing strategy. For two of the three product groups, the self-stated approach provided significantly inflated measures of heterogeneity for the individual products and bundles. Hence this method led to excessively high "optimal" prices for the individual products and the bundle. However, for the third group, the direction of bias in the heterogeneity of reservation prices varied across products and the bundle. These results show that the effect of the self-stated method on prices cannot be specified a priori.

Several areas remain open for future research. Our model should be extended to allow for more general classes of utility function than the quasilinear. Our validity tests are based on choice predictions for a holdout sample. Future research should further assess the validity of our approach by using the estimated reservation prices to predict actual consumer choice. Studies are also necessary to test model robustness for other product classes and new products, in particular. The optimization algorithm implicitly assumes that the firm has a single-period horizon. Future research should extend the optimization algorithm by allowing the firm to behave strategically by choosing product-line pricing strategy to influence the market growth rate and the future joint distribution of reservation prices for the individual products and the bundles.

Acknowledgments

The authors thank the former and present Editors, the Area Editor, and three anonymous reviewers for their constructive suggestions and Asim Ansari, Eric Anderson, and Surendra Rajiv for their helpful comments. Kamel Jedidi acknowledges the research support of Columbia Business School, Columbia University. Sharan Jagpal acknowledges the research support of the Faculty of Management, Rutgers University, and Puneet Manchanda acknowledges the research support from the True North Communications Inc. Faculty Research Fund and the Kilts Center for Marketing at the Graduate School of Business, University of Chicago.

Appendix: Reservation Prices and Utility Theory

Proposition. *Maximizing surplus and maximizing utility are equivalent provided the utility function is quasilinear.*

For simplicity, consider any set of two products and the corresponding bundle $\Phi = \{\phi_1, \phi_2, \phi_3\}$ with corresponding prices $P = \{P_1, P_2, P_3\}$ where "1" and "2" index the individual products and "3" indexes the bundle. Consider any individual $i\,(i = 1, \ldots, I)$ with income w_i and utility function $U_i(\phi_m, z_i)$, where z_i denotes the composite good consisting of all other goods the consumer purchases, measured in some individual-specific basket. Let p_i^w denote the price of the composite good. Assume that individual i consumes only one unit of product/bundle ϕ_m, $m = 1, 2, 3$. Then, the budget constraint facing individual i is $p_i^w z_i + P_m = w_i$. Hence, individual i's indirect utility function is $U_i(\phi_m, (w_i - P_m)/p_i^w)$ if product/bundle ϕ_m is purchased and $U_i(0, w_i/p_i^w)$ if the individual chooses the no-purchase option.

Suppose the utility function is quasilinear. Then the indirect utility function for individual i when product/bundle ϕ_m is purchased is

$$U_i\left(\phi_m, \frac{w_i - P_m}{p_i^w}\right) = u_{im} + \alpha_i \frac{w_i - P_m}{p_i^w}, \tag{A.1}$$

where u_{im} denotes the utility of product or bundle m and $\alpha_i > 0$ is the marginal utility of income. Note that Equation (A.1) assumes that for any individual i, α_i is constant over products and bundles; however, α_i varies across individuals. If none of the products/bundles is purchased, the indirect utility function reduces to

$$U_i\left(0, \frac{w_i}{p_i^w}\right) = \alpha_i \frac{w_i}{p_i^w}. \tag{A.2}$$

Faced with the choice set Φ, individual i will purchase product/bundle ϕ_m iff $U_i(\phi_m, (w_i - P_m)/p_i^w) \geq U_i(0, w_i/p_i^w)$ and has the maximum utility in the choice set $\mathbf{\Phi}$. Let $\beta_i = \frac{\alpha_i}{p_i^w}$. Then, individual i will purchase $\phi_m (m = 1, 2, 3)$ if

$$u_{im} - \beta_i P_m = \max\{u_{i1} - \beta_i P_1, u_{i2} - \beta_i P_2, u_{i3} - \beta_i P_3, 0\}, \tag{A.3}$$

and will choose the no-purchase option if

$$\max\{u_{i1} - \beta_i P_1, u_{i2} - \beta_i P_2, u_{i3} - \beta_i P_3\} < 0. \tag{A.4}$$

Note that the term $\alpha_i(w_i/p_i^w)$ is irrelevant because it does not vary across choice alternatives.

A consumer's reservation price for a product/bundle is the price at which s/he is indifferent to buying and not buying the product, given the consumption alternatives available to the consumer (see Jedidi and Zhang 2002, pp. 1352–1353). Let r_{im} be individual i's reservation price for product/bundle ϕ_m. Then, by definition,

$$U_i\left(\phi_m, \frac{w_i - r_{im}}{p_i^w}\right) - U_i\left(0, \frac{w_i}{p_i^w}\right) \equiv 0. \tag{A.5}$$

Using Equation (A.1) we see that individual i's reservation price for product/bundle ϕ_m is defined by

$$r_{im} = \frac{u_{im}}{\beta_i}. \tag{A.6}$$

Dividing each of the terms in Equations (A.3) and (A.4) by β_i, we see that individual i will purchase $\phi_m (m = 1, 2, 3)$ if

$$r_{im} - P_m = \max\{r_{i1} - P_1, r_{i2} - P_2, r_{i3} - P_3, 0\}, \tag{A.7}$$

and chooses the no-purchase option if

$$\max\{r_{i1} - P_1, r_{i2} - P_2, r_{i3} - P_3\} < 0. \tag{A.8}$$

However, $s_{im} = r_{im} - P_m$ is the surplus derived from purchasing product/bundle ϕ_m. Hence, maximizing utility is equivalent to maximizing surplus.

References

Adams, William J., Janet L. Yellen. 1976. Commodity bundling and the burden of monopoly. *Quart. J. Econom.* **90** 475–498.

Albert, J., S. Chib. 1993. Bayesian analysis of binary and polychotomous response data. *J. Amer. Statist. Assoc.* **88** 669–679.

Allenby, Greg M., Peter E. Rossi. 1998. Marketing models of heterogeneity. *J. Econometrics* **89** 57–78.

Anderson, Simon P., André de Palma, Jacques-François Thisse. 1992. *Discrete Choice Theory of Product Differentiation*. MIT Press, Cambridge, MA.

——, Luc Leruth. 1993. Why firms may prefer not to price discriminate via mixed bundling. *Internat. J. Industrial Organization* **11** 49–61.

Ansari, Asim, S. Siddarth, Charles B. Weinberg. 1996. Pricing a bundle of products or services: The case of nonprofits. *J. Marketing Res.* **33** 86–93.

Bakos, Yannis, Erik Brynjolfsson. 2000. Bundling and competition on the internet. *Marketing Sci.* **19** 63–82.

Becker, Gordon M., Morris H. DeGroot, Jacob Marschak. 1964. Measuring utility by a single-response sequential method. *Behavioral Sci.* **9** 229–232.

Carlin, Bradley P., Thomas A. Louis. 1996. *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman and Hall, London.

Chib, S., E. Greenberg. 1995. Understanding the Metropolis-Hastings algorithm. *Amer. Statist.* **49** 327–335.

Desai, Preyas S. (2001). Quality segmentation in spatial markets: When does cannibalization affect product line design? *Marketing Sci.* **20**(3) 265–283.

Economides, Nicholas. (1993). Mixed bundling in duopoly. Working Paper, New York University, Stern School of Business, EC-93–29.

Gabor, Andre, C. W. J. Granger. 1965. The pricing of new products. *Sci. Bus.* **10** 141–150.

Gelman, A., X. L. Meng, H. S. Stern. 1996. Posterior predictive assessment of model fitness via realized discrepancies (with discussion). *Statistica Sinica* **6** 733–807.

Geman, S., D. Geman. 1984. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Machine Intelligence* **6** 721–741.

Goldberg, Stephen M., Paul E. Green, Yoram Wind. 1984. Conjoint analysis of price premiums for hotel amenities. *J. Business* **57** S111–S132.

Hanson, Ward A., R. Kipp Martin. 1990. Optimal bundle pricing. *Management Sci.* **36** 155–174.

Jedidi, Kamel, John Zhang. 2002. Augmenting conjoint analysis to estimate consumer reservation prices. *Management Sci.* **48** 1350–1368.

Kagel, John H. (1995). Auctions: a survey of experimental research. John H. Kagel, A. E. Roth, eds. *The Handbook of Experimental Economics*. Princeton University Press, Princeton, NJ. 501–585.

——, Ronald M. Harstad, Dan Levin. 1987. Information impact and allocation rules in auctions with affiliated private values: A laboratory study. *Econometrica* **55** 1275–1304.

Kirk, Roger E. 1982. *Experimental Design: Procedures for The Behavioral Sciences*. Brooks/Cole Publishing Company, Pacific Grove, CA.

Kopalle, Praveen K., Aradhana Krishna, Joao L. Assuncao. 1999. The role of market expansion on equilibrium bundling strategies. *Managerial Decision Econom.* **20** 365–377.

Lewbel, A. 1985. Bundling of substitutes or complements. *Internat. J. Industrial Organization* **3** 101–107.

Louviere, J. J., G. Woodworth. 1983. Design and analysis of simulated consumer choice on allocation experiments: An approach based on aggregate data. *J. Marketing Res.* **20** 350–367.

McAfee, R. P., J. McMillan, M. D. Whinston. 1989. Multiproduct monopoly, commodity bundling, and correlation of values. *Quart. J. Econ.* **114** 371–384.

Milligan, Glenn W. 1980. An examination of the effect of six types of error perturbation on fifteen clustering algorithms. *Psychometrika* **45** 325–342.

Monroe, K. B. 1990. *Pricing: Making Profitable Decisions*. McGraw-Hill, New York.

Noble, Peter M., Thomas S. Gruca. 1999. Industrial pricing: Theory and managerial practice. *Marketing Sci.* **18**(3) 435–454.

Salinger, M. A. 1995. A graphical analysis of bundling. *J. Business* **68** 85–98.

Schmalensee, Richard. 1984. Gaussian demand and commodity bundling. *J. Business* **57** 211–230.

Thaler, Richard. 1985. Mental accounting and consumer choice. *Marketing Sci.* **4** 199–214.

Venkatesh, R., Wagner A. Kamakura. 2003. Optimal bundling and pricing under a monopoly: Contrasting complements and substitutes from independently valued products. *J. Business*. Forthcoming.

——, Vijay Mahajan. 1993. A probabilistic approach to pricing a bundle of products or services. *J. Marketing Res.* **30** 494–508.

Vickrey, William. 1961. Counter speculation, auctions, and competitive sealed tenders. *J. Finance* **16** 8–37.

Villas-Boas, J. Miguel. (1998). Product line design for a distribution channel. *Marketing Sci.* **17**(2) 156–169.

Wertenbroch, Klaus, Bernd Skiera. 2000. Measuring consumers' willingness to pay at the point of purchase. *J. Marketing Res.* **34** 228–241.

Wuebeker, Georg, Vijay Mahajan. 1999. A conjoint analysis-based procedure to measure reservation price and to optimally price product bundles. *Optimal Bundling: Marketing Strategies for Improving Economic Performance*. Ralph Fuerderer, Andreas Herrmann, Georg Wuebeker, eds. Springer-Verlag, Berlin, 157–174.

This paper was received July 31, 2000, and was with the authors 10 months for 3 revisions; processed by Chuck Weinberg.