MARKETING SCIENCE
Vol. 23, No. 3, Summer 2004, pp. 391–406
ISSN 0732-2399 | EISSN 1526-548X | 04 | 2303 | 0391

DOI 10.1287/mksc.1030.0032


# A Choice Model with Conjunctive, Disjunctive, and Compensatory Screening Rules

Timothy J. Gilbride
Mendoza College of Business, University of Notre Dame, Notre Dame, Indiana 46556, Timothy.J.Gilbride.2@nd.edu

Greg M. Allenby
Fisher College of Business, 2100 Neil Avenue, Ohio State University, Columbus, Ohio 43210, allenby.1@osu.edu

Many theories of consumer behavior involve thresholds and discontinuities. In this paper, we investigate consumers' use of screening rules as part of a discrete-choice model. Alternatives that pass the screen are evaluated in a manner consistent with random utility theory; alternatives that do not pass the screen have a zero probability of being chosen. The proposed model accommodates conjunctive, disjunctive, and compensatory screening rules. We estimate a model that reflects a discontinuous decision process by employing the Bayesian technique of data augmentation and using Markov-chain Monte Carlo methods to integrate over the parameter space. The approach has minimal information requirements and can handle a large number of choice alternatives. The method is illustrated using a conjoint study of cameras. The results indicate that 92% of respondents screen alternatives on one or more attributes.

*Key words*: conjoint analysis; noncompensatory decision process; hierarchical Bayes; revealed choice; attribute screening; consideration sets; elimination by aspects
*History*: This paper was received October 15, 2002, and was with the authors 1 month for 1 revision; processed by Kannan Srinivasan.

## 1. Introduction

Many theories of consumer behavior involve thresholds and discontinuities. The amount consumers are willing to pay for a lottery ticket is different than the amount they will pay to insure against a commensurate loss. A threshold occurs at the point where a gain turns into a loss. Consumers in the market for a new minivan may not consider offerings above a certain price because they exceed their budget constraints, or because past experience has taught them that higher-priced vehicles are not a good value. When going out for dinner, the restaurants considered may be very different than those considered at lunch. Prospect theory, microeconomics, reference prices, and consideration sets all involve abrupt behavior changes under certain discrete circumstances.

Thresholds and discontinuities are problematic for empirical researchers because they can lead to a likelihood surface that is not differentiable. In a discrete-choice model, the presence of a consideration set leads to abrupt changes in the choice probability as items move into and out of the set. Gradient-based estimation procedures, such as Newton-Raphson and the method of scoring, are not appropriate in these situations because they require a continuous and differentiable likelihood function. Researchers have used various methods to "smooth out" the likelihood by introducing additional parameters and error terms into the model. We propose an alternative approach that deals directly with the irregularity of the likelihood surface, preserving the thresholds and discontinuities that are implied by theory.

In this paper we investigate consumers' use of screening rules as part of a discrete-choice model. Alternatives that pass the screen are evaluated in a manner consistent with random utility theory; alternatives that do not pass the screen have a zero probability of being chosen. Behaviorally, consumers adopt a "rule" to screen alternatives; this rule may be the result of previous learning, information-processing constraints, or solving some previous constrained optimization problem. The reduced-form model we propose is consistent with various theories presented in the consideration set and information-processing literature.

The proposed model accommodates conjunctive, disjunctive, and compensatory screening rules. It is formulated for revealed choice data when process measures or other data, such as acceptable alternatives or attributes, are not available. A primary contribution of this paper is the formulation and estimation of a choice model that reflects a discontinuous decision process. Our approach employs the Bayesian technique of data augmentation and uses Markov-chain Monte Carlo methods to integrate over the parameter space. The approach has minimal information requirements and can handle a large number of choice alternatives.

The organization of the paper is as follows. In §§2 and 3 we discuss two-stage decision processes and how they relate to screening rules. In §4 we describe our modeling approach and how it overcomes difficulties associated with a discontinuous likelihood surface. Section 5 introduces an empirical example in the context of a choice-based conjoint study for a new camera type. Results are presented in §6, where we find that 92% of the survey respondents use some form of noncompensatory screening when choosing among the choice options. Implications and directions for future research are discussed in §7.

## 2. Two-Stage Decision Processes

Consumers use various decision rules to simplify complicated decision tasks. The choice model proposed in this paper is consistent with two-stage decision processes wherein a subset of product alternatives is selected from the universal set, and the final product choice is from the reduced set. Theoretical motivation for this two-stage process can be found in the consideration set and information-processing literatures.

The term "consideration set" has been used in various ways in the marketing literature. Shocker et al. (1991, p. 183) define consideration sets as being "purposefully constructed and can be viewed as consisting of those goal-satisfying alternatives salient or accessible on a particular occasion." Hence, consideration sets are linked to usage occasion. Alternatively, Hauser and Wernerfelt (1990) define consideration sets as a subset of offerings that receive serious consideration during the purchase occasion. Regardless of its use, researchers in marketing have generally operationalized consideration sets as comprising alternatives that survive a screening process (see Häubl and Trifts 2000).

Consideration sets have been modeled as the solution to a consumer optimization problem. Hauser and Wernerfelt (1990) provide a model where the expected utility of consumption must exceed the cognitive cost of evaluation for an alternative to be considered. Roberts and Lattin (1991) operationalize this approach using survey data, and Mehta et al. (2003) extend this structural approach to scanner panel data where explicit information on consideration sets is not available. These approaches assume that the overall utility of the offering determines inclusion in the consideration set (see also Montgomery et al. 2003).

An information-processing view of the consumer leads to consideration sets determined by product attributes, not overall utility. These models typically have a less effortful first stage, followed by a more comprehensive second stage. Payne (1976) and Bettman and Park (1980) provide evidence that the formation of the consideration set is linked to a subset of attributes and that the final selection is more holistic. Bettman et al. (1998) provide an overview of consumer decision strategies, including the use of various conjunctive (i.e., intersection) and disjunctive (i.e., union) rules that can be used for set formation. Bettman (1979), Shugan (1980), and Bettman et al. (1990) show that adopting a screening rule based on a subset of attributes is a rational outcome for a decision maker willing to tradeoff accuracy against cognitive effort.

Our model relies entirely on revealed choices and can accommodate holistic (i.e., overall utility) and attribute-based screens. It does not incorporate information that can be used to determine the process by which "seriously considered" alternatives are identified, nor does it require respondents to identify the set of alternatives that pass the screen. Hence, it is a reduced-form model that is consistent with the literature on consideration sets. To minimize confusion with prior research, we refer to the set of alternatives that pass the screening rule in a given context as the choice set.

## 3. Screening Rules

The screening rules restrict the set of alternatives that are evaluated for final selection. The alternatives included in the choice set are identified with an indicator function, $\mathrm{I}(x_j, \gamma)$, that equals one if the decision rule is satisfied and equals zero otherwise. The choice model can then be written in general form as

$$\Pr(i) = \Pr\big(V_i + \varepsilon_i > V_j + \varepsilon_j \text{ for all } j \text{ such that } \mathrm{I}(x_j, \gamma) = 1\big) \tag{1}$$

where $V_i$ denotes the deterministic portion of utility of choice alternative $i$, $\varepsilon_i$ is the stochastic portion and it is assumed that the stochastic terms are distributed independently, $x_j$ denotes a generic argument of an indicator function (I) that reflects the decision rule applied to the $j$th offering. Alternative screening rules can be investigated with the indicator function relationship, $\mathrm{I}(x_j, \gamma) = 1$. Flexibility exists in the specification of this relationship because its complexity does not increase the complexity of estimating the choice model, as discussed in §4.

For example, if price is used to screen the alternatives so that only offerings above a lower threshold are considered, then $\mathrm{I}(x_j, \gamma) = \mathrm{I}(\text{price}_j > \gamma)$. If an upper threshold for price is present, then negative price can be used as an argument, $\mathrm{I}(x_j, \gamma) =$

$I(-\text{price}_j > \gamma)$. Here, $\gamma$ is a parameter to be estimated. This formulation allows for dynamic choice sets when an attribute (e.g., price) changes over the choice history.

A compensatory screening rule is one where the deterministic portion of the utility of the offering must exceed a threshold value to be acceptable.

$$\text{Compensatory Rule:} \quad I(V_j > \gamma) = 1. \tag{2}$$

This screening rule is consistent with consideration sets based on consumer optimization models (see especially Roberts and Lattin 1991). Assuming that the stochastic portion of utility has an expected value of 0, $V_j$ is the expected utility from the offering and $\gamma$ the cognitive cost of considering an offering.

Two noncompensatory screening rules consistent with the information-processing literature are also modeled: conjunctive and disjunctive. Choice sets formed by a conjunctive decision rule require that an alternative be acceptable on all relevant attributes for it to be included. A conjunctive rule is formed by multiplying indicator functions across the attributes ($m$) of an offering:

$$\text{Conjunctive Rule:} \quad \prod_m I(x_{jm} > \gamma_m) = 1, \tag{3}$$

where $x_{jm}$ is the level of attribute $m$ for choice alternative $j$. The cutoff value, $\gamma_m$, is the smallest level of the attribute that needs to be present for the decision maker to consider the offering. If the cutoff value is smaller than all levels of the attribute, then the attribute is not used to screen. For nominally scaled attributes (i.e., 0, 1 attributes), a cutoff value greater than zero indicates that the attribute is required. Use of the product $\prod_m$, results in a conjunctive combination of the factors where all components must be satisfied for the combination to be satisfied.

The conjunctive screening rule is consistent with an elimination-by-aspects screening process. One might hypothesize that a consumer identifies her most important attribute and eliminates any alternatives not exceeding a certain threshold. Then she may move onto her second-most important attribute, eliminating any alternatives not exceeding a certain level on that attribute, etc. Note that the order in which the attributes are examined doesn't matter: The end result is that each attribute must be above a specific threshold to be considered. This is identical to the conjunctive screening rule used in the model.

A disjunctive decision rule is one where at least one of the attribute levels is acceptable:

$$\text{Disjunctive Rule:} \quad \sum_m I(x_{jm} > \gamma_m) \geq 1. \tag{4}$$

We note that the conjunctive and disjunctive screening rules directly relate the choice set to specific levels of the attributes. Estimates of $\gamma_m$ will indicate which attributes, and what levels, are critical to consumers.

Screening rules consistent with consumer optimization (compensatory) and information-processing (conjunctive and disjunctive) theories have been proposed to represent two-stage decision making. Our reduced-form model, Equation (1), however, does not require that consumers form a choice set first, and then choose from that set. The mathematical representation is silent on any specific procedure. However, it does capture the sharp discontinuity in the underlying behavioral model: Alternatives with unacceptable attribute levels (or combinations of attributes) have a zero probability of selection. A method for estimating the model is discussed next.

## 4. Estimating Choice Models with Screening Rules

Consideration or choice-set models have been applied in a variety of domains, including the analysis of survey data (Gensch 1987), scanner panel data (Manrai and Andrews 1998), and data collected in laboratory settings (Johnson and Meyer 1984). Johnson et al. (1989) and Andrews and Manrai (1998) demonstrate that parameter bias and incorrect inferences result when the decision process is not correctly modeled.

Choice sets induce a complicated structure in the likelihood function. Consider, for example, a simple discrete-choice model without choice sets. The choice probability is

$$\begin{aligned} \Pr(i) &= \Pr(V_i + \varepsilon_i > V_j + \varepsilon_j \text{ for all } j) \\ &= \Pr(\varepsilon_j < V_i - V_j + \varepsilon_i \text{ for all } j) \\ &= \int_{-\infty}^{+\infty} \left[ \int_{-\infty}^{V_i - V_j + \varepsilon_i} \cdots \int_{-\infty}^{V_i - V_m + \varepsilon_i} f(\varepsilon_j) \cdots \right. \\ &\qquad \left. f(\varepsilon_m)\, d\varepsilon_j \cdots d\varepsilon_m \right] f(\varepsilon_i)\, d\varepsilon_i, \end{aligned} \tag{5}$$

where $V_i$ denotes the deterministic portion of utility of choice alternative $i$, $\varepsilon_i$ is the stochastic portion, and it is assumed that the stochastic terms are distributed independently. The probit model is obtained by assuming that the stochastic terms are distributed normal, and the logit model is obtained when the stochastic terms are distributed extreme value. The discrete-choice probabilities are derived by partitioning the domain of the error space corresponding to a choice outcome, and integrating the error densities over the corresponding region. The regions for a simple discrete-choice model are indicated in the upper and lower limits of the integrals.

The likelihood for an observed sequence of choices is the product of the choice probabilities:

$$\ell(V \mid data) \equiv \pi(data \mid V) = \prod_i \Pr(i). \tag{6}$$

The likelihood surface for simple discrete-choice models is continuous and globally concave (Maddala 1983). Small changes in the deterministic portion of utility ($V$) result in small changes in the limits of integration and small changes in the resulting choice probabilities. Moreover, unique values of the deterministic portion exist that maximize the probability of the observed choices. The regularity of the likelihood surface allows use of gradient-based methods (e.g., Newton-Raphson) to solve for values of the deterministic portion ($V$) that maximize the likelihood.

The presence of choice sets destroys the regularity of the likelihood function. As choice alternatives enter and leave the set, the choice probabilities are obtained by integrating over an error space of changing support. Adding an alternative ($k$) to the choice set results in an additional factor in Equation (5), with an additional stochastic error ($\varepsilon_k$) and appropriate limits of integration. The likelihood surface for a choice-set model is therefore discontinuous and not necessarily globally concave.

Researchers have dealt with the nonregularity of the likelihood surface by writing the probability as the product of a (conditional) second-stage probability and a marginal choice-set probability:

$$\Pr(i, C) = \Pr(i \mid C) \times \Pr(C), \tag{7}$$

where $C$ denotes the choice set that reduces the number of alternatives in Equation (5). The likelihood function is "smoothed out" by summing over choice sets:

$$\Pr(i) = \sum_{C} \Pr(i \mid C) \times \Pr(C). \tag{8}$$

This approach, coupled with the assumption that the values of the deterministic components $\{V_i\}$ are constant across the choice sets, results in a smooth and globally concave likelihood surface.

The estimation of choice-set models using Equation (8) has historically been approached sequentially, sometimes with independent data. Jedidi et al. (1996) and Roberts and Lattin (1991) have respondents indicate the acceptability of alternatives in a survey. Ben-Akiva and Boccara (1995) also use survey data and employ a logit model for the choice-set probability. When independent data are available to identify the acceptability of an offering, the evaluation of choice-set models can employ standard maximum-likelihood techniques.

If independent data are not available to identify the choice set, it becomes a latent construct that must be simultaneously estimated with the observed choices. Complications arise in the evaluation of the choice probabilities because the partition of the error space described in Equation (5) must also be partitioned with respect to the choice set (see Mehta et al. 2003). Only one realization of the error vector, $\varepsilon$, is associated with each observation, and it is not appropriate to estimate the two components in Equation (7) by assuming that each are associated with a different realization.

An alternative to partitioning the error space is to parameterize and estimate the point mass of each choice set in Equation (8). Andrews and Srinivasan (1995) use this approach on scanner panel data, where the probability for each of the $2^J - 1$ possible choice sets is enumerated. Bronnenberg and Vanhonacker (1996) propose a different likelihood function where only $J$ (one for each brand) additional probabilities must be computed. Each of these models employs a compensatory function to determine choice-set probability.

One method of dealing with the dimensionality of the problem is to place a priori limits on the allowable choice sets. Fader and McAlister (1990) assume the choice set is either the set of previously purchased brands currently on promotion or all previously purchased brands, and they estimate the probability of each being the actual choice set. Siddarth et al. (1995) also restrict the number of choice sets to two: the universal set or a subset of brands stochastically determined based on past purchases.

Chiang et al. (1999) consider Equation (7) as a hierarchical Bayes model, where the $2^J - 1$ sets are assigned a prior distribution and posterior estimates of set composition are obtained. While this approach avoids the complications involved with identifying suitable partitions of the error space to derive probabilities, the choice sets are not linked to the characteristics of the offerings other than through the fact that the chosen brand is an element of the set. Choice sets with high posterior probabilities comprise offerings that have been chosen and exclude offerings that have not been chosen.

By contrast, our reduced-form model, Equation (1), relies on screening rules and not on the enumeration of choice sets. The size of the universal set is not a concern, nor is it necessary to specify a priori which attributes are being used to screen (e.g., price in Mehta et al. 2003 or promotion in Fader and McAlister 1990) or place limits on the number of the choice sets. Swait (2001) proposes a different reduced-form model that is structurally a compensatory model but mimics various decision rules depending on the actual parameter estimates. A strength of Swait's (2001) model is that it can be estimated using standard maximum-likelihood techniques. However, it requires individuals to identify "unacceptable" levels of attributes, and it cannot estimate the prevalence of screening in the sample.

Our model structure allows for screening alternatives based on the attribute levels and/or the overall value of the offering. Attribute level screening can employ either conjunctive or disjunctive rules. We navigate the complicated partitioning of the error space by employing data augmentation and Markov chain Monte Carlo methods (MCMC). We next describe the use of data augmentation to evaluate the standard probit model in Equation (5) above. We then show how this method can be used to estimate choice models with screening rules.

**Data Augmentation**
Direct evaluation of the probability in Equation (5) can be avoided in a Bayesian approach to estimation if the parameters $\{V\}$ are augmented with a vector of latent variables, $z$ (see Tanner and Wong 1987), where

$$z_i = V_i + \varepsilon_i, \quad \varepsilon_i \sim \text{Normal}(0, 1), \tag{9}$$

and the model is written hierarchically:

$$y \mid z, \tag{10a}$$

$$z \mid V, \tag{10b}$$

where $y$ denotes the observed choice outcomes. Equation (10a) reflects the censoring of the continuous variables $z$, where only the offering with the maximum value of $z$ is revealed to the researcher by the choice data. Equation (10b) reflects the assumption that the latent variables are normally distributed, with the mean of each element equal to $V_i$ and variance one. In marketing applications of the probit model, the mean is typically parameterized $V_i = x_i'\beta$. Markov chain Monte Carlo estimation of the probit model proceeds by drawing iteratively from the two conditional distributions:

$$\pi(z \mid y, V) = \text{Truncated Normal}(V, \mathbb{1}), \tag{11a}$$

$$\pi(V \mid z) = \text{Normal}(\bar{z}, \mathbb{1}/n), \tag{11b}$$

where "$\mathbb{1}$" is the identity matrix, and the truncation in Equation (11a) corresponds to the alternative selected being the maximum draw. If $V_i$ is parameterized as $x_i'\beta$, then the mean and variance in Equation (11b) would correspond to the OLS regression estimate and covariance matrix. Detailed treatments of estimating the probit model with data augmentation can be found in Albert and Chib (1993), McCulloch and Rossi (1994), Allenby and Rossi (1999), and Edwards and Allenby (2003).

The upper limits of the integral in Equation (5) that complicate the evaluation of the choice probability are replaced with a much simpler procedure that involves drawing from a truncated normal distribution. The truncation points correspond to the limits of integration, and the Markov chain is used to navigate the domain of the error space consistent with the observed data to arrive at the posterior distribution of model parameters. Thus, data augmentation provides a clever way of avoiding the estimation of the integral.

**Application to Choice Models with Screening Rules**
If an alternative is in the choice set, then its choice probability is determined relative to the other offerings in the set. If an alternative does not satisfy the screening rule, then its choice probability is zero. The choice model can be written hierarchically as

$$y \mid z, \text{I}(x, \gamma), \tag{12a}$$

$$z \mid V, \tag{12b}$$

where $z_i = V_i + \varepsilon_i$, and estimation proceeds by drawing iteratively from the conditional distributions:

$$z \mid y, V, \text{I}(x, \gamma), \tag{13a}$$

$$\gamma \mid y, z, x, \tag{13b}$$

$$V \mid z. \tag{13c}$$

The conditional distribution in Equation (13a) is a truncated normal distribution for those items in the choice set. That is, the element of $z$ for the chosen alternative is greater than the elements of $z$ that correspond to the alternatives not chosen, given that they are in the choice set. If an alternative is not in the choice set, which is determined by the indicator function $\text{I}(x, \gamma)$, then $z$ is drawn from a nontruncated distribution. Here we assume that $\text{I}(x, \gamma)$ does not involve the deterministic component of utility, $V$; changes to the hierarchy to accommodate this screening rule are discussed below.

The conditional distribution for the cutoff parameters, $\gamma$, in Equation (13b) is dependent on the observed data, $y$, and the augmented parameter, $z$. The conditional distribution takes the form of an indicator function because Equation (12a) is a deterministic function that censors the continuous latent variables $z$, identifying the considered offering with the highest value. Variation in the value of the cutoff parameter changes the set of considered alternatives. Permissible values of the cutoff parameters are those that lead to a choice set where the maximum of the augmented variable, $z$, corresponds to the observed choice. If a value of the cutoff leads to a choice set where the alternative with the maximum $z$ is not the alternative chosen, then the cutoff parameter is not permissible. This can happen because some values of the augmented parameter in Equation (13a) are drawn from a truncated distribution while others are not. Variation in $z$ results in different choice sets by defining the permissible values of the cutoffs.

Finally, the conditional distribution in Equation (13c) is identical to Equation (11b). Standard normal distribution theory can be used to generate the draw of $V$. If $V$ is parameterized as a regression function, $V = X\beta$, then draws of $\beta$ are generated from a normal distribution with mean and covariance matrix corresponding to the standard OLS model. Details of the estimation routine for our empirical study, reported below, are provided in the appendix.

The Markov chain defined by Equations (13a)–(13c) generates draws from the full conditional distribution of model parameters despite the irregularity of the likelihood surface. Given the cutoffs, $\gamma$, the model (Equation (1)) becomes a standard discrete-choice model. Given the set of augmented variables, $\{z\}$, some values of the cutoff are acceptable and others are not. Across multiple draws of the augmented variables, however, the acceptable range of $\gamma$ varies, and the Markov chain converges in distribution to the full conditional distribution of all model parameters. The augmented variable $z$ reflects a single error term that gives rise to the joint probability of the consideration set and the resulting choice.

Recall that our model does not rely on process data or respondents indicating which alternatives or which attribute levels are acceptable. Intuitively, the logic for identifying acceptable levels of the attributes relies on the revealed choices. If a respondent chooses an alternative, then the screening rule must be satisfied. The presence of attributes that are common to more than one choice alternative allows identification of the cutoff parameter, $\gamma$. Across all alternatives and choice occasions, there is a set of $z$s, $V$s, and $\gamma$s that are consistent with the revealed choices. We use the Markov chain to explore the posterior distribution of the allowable values of these parameters.

The compensatory decision rule requires a change in the hierarchical model. Because the deterministic portion of utility, $V$, is used to determine both the choice set and the final choice, the hierarchy is

$$z \mid y, V, \mathrm{I}(V, \gamma), \tag{13a$'$}$$

$$\gamma \mid y, z, V, \tag{13b$'$}$$

$$V \mid z, y. \tag{13c$'$}$$

The exact form of these conditional distributions and the estimation algorithm is contained in the Appendix.

In addition to the form of the screening rule, researchers can specify that the cutoff parameters ($\gamma$) take on a limited range of values. If an attribute, $x$, is continuous (e.g., price), then $\gamma$ can also be continuous. However, if the attribute takes on a restricted number of values, then the cutoff is not well identified over an unrestricted range. In a conjoint analysis, for example, there are rarely more than four or five levels of an attribute. In these cases one can assume that the cutoff parameter takes on a small number of finite values by a prior assumption that $\gamma$ is distributed multinomial over a grid of value and estimate the posterior distribution of the multinomial mass probabilities. Next, we explain this approach in the context of our empirical application.

## 5. Data and Models

The screening rule model is illustrated in the context of a discrete-choice conjoint study of consumer preferences for the features of a new camera format, the advanced photo system (APS). This photographic system offers several potential advantages over standard 35-mm film systems, including midroll film change, camera operation feedback, and a magnetic layer on the film that offers the ability to annotate the picture with titles and dates. These data have previously been described in Allenby et al. (2002).

Respondents were screened and recruited by telephone and interviewed at field locations in nine metropolitan areas. A total of 302 respondents participated. The screen was designed to exclude individuals who had little or no interest in photography, and retained approximately 55% of the population contacted. This compares to approximately 67% of the population that owns still cameras (Simmons 1997). Individuals participating in the study were therefore assumed to be good prospects for cameras in the price range of the new APS offerings and representative of the APS camera market.

Respondents were asked to evaluate 14 buying scenarios. The order of the buying scenarios was randomized. Each scenario contained full profile descriptions of three compact 35-mm cameras, three new APS cameras, and a seventh no-buy option. Table 1 lists the attributes that describe each of the cameras and identifies the attributes that are new and specific to APS cameras. The order of the attributes was the same in each scenario. Table 1 also displays restrictions on the availability of certain attributes. For example, the $4x$ zoom lens was available only on the "high" body style. Prices ranged from \$41 to \$499, the average price for APS cameras was \$247, and the average price of the compact 35-mm cameras was \$185.

Each buying scenario requires the respondent to engage in a complex decision. With six offerings described on eight attributes, plus the no-buy option, we consider this to be a good dataset with which to investigate the use of screening rules. The no-buy option was selected 14.7% of the time or an average of 2.1 times in the 14 choice tasks. One of the APS camera offerings was chosen 39.9% of the time, and one of the 35-mm cameras was selected 45.4% of the time. In the analysis that follows, 12 choice tasks per respondent were used to calibrate the model and the final two retained for predictive testing. This results in 3,624 observations for model estimation and 604 for hold-out validation.

**Table 1 Camera Attributes and Levels**

| Attribute/level | Available in "low" body style | Available in "medium" body style | Available in "high" body style |
|---|---|---|---|
| A. Basic body style and standard features | | | |
| A1. Low | × | | |
| A2. Medium | | × | |
| A3. High | | | × |
| B. Midroll change* | | | |
| B1. None | × | × | × |
| B2. Manual | × | | |
| B3. Automatic | × | × | × |
| C. Annotation* | | | |
| C1. None | × | × | × |
| C2. Preset list | × | × | × |
| C3. Customized list | × | × | × |
| C4. Custom input method 1 | | | × |
| C5. Custom input method 2 | | × | × |
| C6. Custom input method 3 | | | × |
| D. Camera operation feedback* | | | |
| D1. No | × | × | × |
| D2. Yes | | | × |
| E. Zoom | | | |
| E1. None | × | × | × |
| E2. 2*x* | | × | × |
| E3. 4*x* | | | × |
| F. Viewfinder | | | |
| F1. Regular | × | × | × |
| F2. Large | | | × |
| G. Camera settings feedback | | | |
| G1. None | × | | |
| G2. LCD | × | × | × |
| G3. Viewfinder | | × | × |
| G4. Both LCD & viewfinder | | × | × |

*Unique to APS camera at the time of the study; all other features available on both APS and 35-mm cameras at the time of the study.

Five models are fit to the data. See Table 2 for a summary of the models. In each model we make the assumption that the no-buy option is always in the choice set. "Not buying" is always an option, both in the empirical conjoint study we use and in real markets. The first model is a heterogeneous probit model that assumes all choice options are considered by the respondent. For identification, effects-level coding is used for the nonprice attributes, with the lowest level of each attribute (except for body style) set to zero. Price is parameterized as a natural logarithm. For the no-buy option, all attributes are set to zero (including the natural logarithm of price). The effects-level coding results in 18 parameters, 17 corresponding to binary indicator variables and the 18th corresponding to ln(price). To fix the scale in the probit model, the error term is assumed to be independent and identically distributed across choice tasks with unit variance. The latent continuous variable in our model setup is therefore

$$\Pr(j)_{hi} = \Pr(z_{hij} > z_{hik} \text{ for all } k), \tag{14}$$

$$z_{hi} = X_{hi}\beta_h + \varepsilon_{hi}, \quad \varepsilon_{hi} \sim \text{Normal}(0, \mathbb{1}), \tag{15}$$

$$\beta_h \sim \text{Normal}(\bar{\beta}, \Sigma_\beta), \tag{16}$$

where $h$ indexes the respondent, $i$ is the index for the buying scenario, and $j$ and $k$ indicate the alternatives within the buying scenario.

The second model assumes that the choice set comprises alternatives with the deterministic value of utility ($V_{hi} = X_{hi}\beta_h$) greater than a threshold value, $\gamma_h$. That is, the selected alternative has greatest utility of the options for which $V_{hik} > \gamma_h$:

$$\Pr(j)_{hi} = \Pr\big(z_{hij} > z_{hik} \text{ for all } k \text{ such that } \mathrm{I}(V_{hik} > \gamma_h) = 1\big). \tag{17}$$

We allow for heterogeneity in the cutoff value, $\gamma_h$ through the normal distribution:

$$\gamma_h \sim \text{Normal}(\bar{\gamma}, \sigma_\gamma^2). \tag{18}$$

A weakness in our modeling approach is that separate "weights" for the screening rule and final selection component of the model are not statistically identified. By identifying a separate error term along with appropriate identification restrictions and using Equation (8), other researchers have allowed for the "consideration utility" to be different than the "selection utility" (cf. Andrews and Srinivasan 1995). While our approach for the compensatory screen is consistent with consumer optimization models, it is less flexible than other models.

The third model assumes that the choice set is formed using a conjunctive screening rule over the $m$ attributes for each alternative:

$$\Pr(j)_{hi} = \Pr\big(z_{hij} > z_{hik} \text{ for all } k \text{ such that } \textstyle\prod_m \mathrm{I}(x_{hikm} > \gamma_{hm}) = 1\big), \tag{19}$$

**Table 2 Model Results**

| Model | In sample LMD* | Predictive Hit** probability | Predictive Hit*** frequency |
|---|---|---|---|
| HB probit | −3,468.8 | 0.391 | 266 |
| Compensatory | −3,449.6 | 0.393 | 267 |
| Conjunctive | −2,990.8 | 0.418 | 276 |
| Disjunctive | −3,586.2 | 0.39 | 267 |
| Structural heterogeniety | −3,012.9 | 0.417 | 277 |

*Log marginal density of the data calculated using the importance-sampling method of Newton and Raftery (1991, p. 21).

**Posterior mean of predicted probability for selected alternative. Calculated from a sample of 300 posterior point estimates.

***Posterior mean of correct predictions for two observations for each respondent; a total of 604 holdout choice sets. Calculated from a sample of 300 posterior point estimates.

where $x_{hikm}$ is the level of the attribute for respondent $h$ in buying scenario $i$ for alternative $k$ and attribute $m$. $\gamma_{hm}$ is a respondent-level parameter representing the threshold, or acceptable level, of attribute $m$ for respondent $h$. We employ two alternative heterogeneity distributions for the threshold parameters. When attribute $m$ is continuously distributed (e.g., price), then we employ a normal distribution to allow for heterogeneity:

$$\gamma_{hm} \sim \text{Normal}(\bar{\gamma}_m, \sigma_m^2). \tag{20}$$

When an attribute takes on a small number of discrete levels, we assume that the threshold parameter is distributed multinomial. For example, consider the zoom lens attribute in Table 1 taking on values of none, 2×, and 4×. We recode the attribute levels (e.g., 0, 1, 2) and specify a grid of possible cutoffs (e.g., −0.5, 0.5, 1.5, 2.5), where the lowest cutoff value indicates that all attribute levels are acceptable, and the highest level indicates that none of the levels is acceptable. For nominally scaled attribute levels where a prior ordering is not obvious (e.g., color), the grid has three points (−0.5, 0.5, 1.5). The full set of attributes, levels, and cutoffs is contained in Table 5. The distribution of heterogeneity of the cutoff parameters with discrete levels is

$$\gamma_{hm} \sim \text{Multinomial}(\theta_m), \tag{21}$$

where $\theta_m$ is the vector of multinomial probabilities associated with the grid for attribute $m$.

The fourth model assumes that the choice set is formed using a disjunctive decision rule:

$$\Pr(j)_{hi} = \Pr\big(z_{hij} > z_{hik} \text{ for all } k \text{ such that } \textstyle\sum_m \text{I}(x_{hikm} > \gamma_{hm}) \geq 1\big). \tag{22}$$

The upper and lower grid points specify the role of the attribute in determining the choice set, depending on the decision rule. An attribute determines membership in the choice set only if it changes the value of the alternative indicator function. For the conjunctive model, if the threshold for attribute $k$ is equal to −0.5, then $\text{I}(x_{hijk} > \gamma_{hk})$ always equals one and the value of $\prod_m \text{I}(x_{hijm} > \gamma_{hm})$ does not change. Because different levels of the attribute do not affect choice sets, we conclude that attribute $k$ is not being used to screen alternatives. For the disjunctive model, if $\gamma_{hk} = 2.5$, then $\text{I}(x_{hijk} > \gamma_{hk})$ always equals zero, and the value of $\sum_m \text{I}(x_{hijm} > \gamma_{hm})$ does not change across different levels of attribute $k$. For the disjunctive model, the upper value of the threshold implies no screening based on that attribute. Figure 1 illustrates the procedure and the interpretation.

The fifth model assumes that consumers are either using a conjunctive (Equation (19)) or a disjunctive (Equation (22)) screening rule. That is, we allow for the presence of structural heterogeneity (Yang and Allenby 2000):

$$\Pr(j)_{hi} = \phi \Pr\big(z_{hij} > z_{hik} \text{ for all } k \text{ such that } \textstyle\prod_m \text{I}(x_{hikm} > \gamma_{hm}) = 1\big) + (1-\phi)\Pr\big(z_{hij} > z_{hik} \text{ for all } k \text{ such that } \textstyle\sum_m \text{I}(x_{hikm} > \gamma_{hm}) \geq 1\big), \tag{23}$$

where $\phi$ is the portion of the sample using the conjunctive decision rule and $(1-\phi)$ is the portion using the disjunctive decision rule. A concise listing of the models is provided in Table 2.

Inference is conducted via Bayesian MCMC methods. The chain converged quickly in all models. Convergence was assessed by inspecting time-series plots from multiple starting points. The chain was run for 10,000 iterations, with the final 5,000 iterations used to estimate the moments of the posterior distributions. Details of the estimation algorithms for all models are provided in the Appendix.

## 6. Results

The results indicate that the conjunctive model fits the data best. In-sample and predictive fits of the five models are presented in Table 2. The log marginal density is computed using the GHK estimator (Keane 1994, Hajivassilious et al. 1996) and the importance sampling method of Newton and Raftery (1994). Hit probability is defined as the average probability associated with the observed choices. Hit frequency is the number of times the predicted choice matches the actual choice. The parameters reported below for the conjunctive model result in 30% of the alternatives screened out of the final choice set, which averaged five (out of seven) alternatives, including the no-buy option.

Fit results for the other models are as follows. The fit for the disjunctive model is slightly worse than the standard probit model, with parameters that lead to only 2% of the alternatives screened from final consideration. The slight decrease in the log marginal density for the disjunctive model is from the increase in the dimension of the model due to the cutoff parameters. Although the compensatory model screened out 23% of the alternatives to form the choice set and has a slightly better in-sample fit than the probit model, its predictive accuracy is nearly identical. The structural heterogeneity model assigned nearly all the respondents (99%) to the conjunctive screening rule model as opposed to the disjunctive model, so the fit statistics closely match those of the conjunctive model.

The results provide evidence that the conjunctive model is superior to both the standard probit and

**Figure 1** **Example for Specifying Cutoffs for Conjunctive and Disjunctive Model Discrete Attributes**

| Attribute | Levels | Recoded Values (x) |
|---|---|---|
| *Zoom Lens* | None | 0 |
| | 2x | 1 |
| | 4x | 2 |

| Possible Cutoffs | Probability of Each Cutoff |
|---|---|
| $\gamma_{(1)} = -0.5$ | $\theta_{(1)}$ |
| $\gamma_{(2)} = 0.5$ | $\theta_{(2)}$ |
| $\gamma_{(3)} = 1.5$ | $\theta_{(3)}$ |
| $\gamma_{(4)} = 2.5$ | $\theta_{(4)}$ |

Scenario #1:
***Screening rule:*** $I(x > \gamma) = 1$
**If $\gamma$ = -0.5**: $I(0 > -0.5) = 1$, $I(1 > -0.5) = 1$, $I(2 > -0.5) = 1$ — *All attribute levels are acceptable*
For the conjunctive model, if $\gamma = -0.5$, then Zoom Lens is not being used to screen alternatives.

Scenario #2:
***Screening rule:*** $I(x > \gamma) = 1$
**If $\gamma$ = 0.5**: $I(0 > 0.5) = 0$, $I(1 > 0.5) = 1$, $I(2 > 0.5) = 1$ — *Camera must have 2x zoom lens to be considered*

Scenario #3:
***Screening rule:*** $I(x > \gamma) = 1$
**If $\gamma$ = 1.5**: $I(0 > 1.5) = 0$, $I(1 > 1.5) = 0$, $I(2 > 1.5) = 1$ — *Camera must have 4x zoom lens to be considered*

Scenario #4:
***Screening rule:*** $I(x > \gamma) = 1$
**If $\gamma$ = 2.5**: $I(0 > 2.5) = 0$, $I(1 > 2.5) = 0$, $I(2 > 2.5) = 0$ — *None of the attribute levels are acceptable*
For the disjunctive model, if $\gamma = 2.5$, then Zoom Lens is not being used to screen alternatives.

nonconjunctive models. The conjunctive model has the largest value of the log marginal density. Despite the addition of 11 cutoff parameters per respondent, the conjunctive model performs as well as or better than more parsimonious models in the hold-out sample. This establishes the predictive validity of the choice model with a conjunctive screening rule.

Tables 3 and 4 report part-worth estimates for the probit and conjunctive models. Table 3 displays the mean and covariance matrix of the part-worths for the probit model, and Table 4 reports the same statistics for the conjunctive model. Substantive differences exist in the part-worth estimates for body style, zoom, camera setting feedback, and price. The conjunctive model estimates are much smaller for these attributes, with the average price coefficient moving to zero. The reason for this difference is discussed below.

The cutoff parameters are reported in Table 5. For discrete attributes, the cutoffs are reported in terms of multinomial point mass probabilities associated with the grid described in Figure 1. The probabilities indicate the fraction of respondents with a cutoff parameter ($\gamma$) that screens out choice alternatives. For the conjunctive model, $\theta_1 = 0.598$ indicates that 59.8% of the respondents are not screening alternatives based on "Body Style," $\theta_2 = 0.324$ implies that 32.4% of the respondents screen out offerings with the low body style, $\theta_3 = 0.059$ implies that 5.9% of the respondents screen out offerings with low and medium body style, and $\theta_4 = 0.015$ implies that 1.5% of the respondents screen on all alternative levels. Estimates of $\theta$ equal to approximately 2% reflect the influence of the prior distribution and can be disregarded for practical purposes.

The use of screening rules to form choice sets is pervasive among the respondents in this research. Analysis of the individual-level cutoff values indicates that 92% of respondents form choice sets using the conjunctive screening rule; 58% are forming choice sets based on a single attribute, 33% are using two attributes, and 2% are using three attributes. A strength of the Bayesian approach outlined in this paper is the ability to make individual-level inferences and estimate not only what attributes are important in forming choice sets, but also who is using screening rules.

Consider, for example, an aggregate analysis of the effect of the price threshold, estimated to be equal to

**Table 3 Part-Worth Estimates for the Probit Model**

| Attribute | Level | Beta | Posterior mean | Attribute | Level | Beta | Posterior mean |
|---|---|---|---|---|---|---|---|
| Body style | Low body | 1 | 0.325 | Camera operation feedback | Operation feedback | 11 | **0.357** |
| | Medium body | 2 | **2.843** | Zoom | 2× zoom | 12 | **1.122** |
| | High body | 3 | **2.224** | | 4× zoom | 13 | **1.477** |
| Midroll change | Manual change | 4 | 0.059 | Viewfinder | Large viewfinder | 14 | **−0.158** |
| | Automatic change | 5 | **0.178** | Settings feedback | LCD | 15 | **0.759** |
| Annotation | Preset list | 6 | **0.362** | | Viewfinder | 16 | **0.757** |
| | Customized list | 7 | **0.678** | | LCD & viewfinder | 17 | **0.902** |
| | Input method 1 | 8 | **−1.036** | Price | ln(price) | 18 | **−0.710** |
| | Input method 2 | 9 | **0.767** | | | | |
| | Input method 3 | 10 | **−0.985** | | | | |

**Covariance Matrix—$\Sigma_\beta$**

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | **19.605** | | | | | | | | | | | | | | | | | |
| 2 | **17.952** | **20.258** | | | | | | | | | | | | | | | | |
| 3 | **15.303** | **17.674** | **17.845** | | | | | | | | | | | | | | | |
| 4 | −0.548 | −0.906 | −1.218 | **1.803** | | | | | | | | | | | | | | |
| 5 | **−2.077** | **−2.232** | **−2.153** | **0.507** | **1.400** | | | | | | | | | | | | | |
| 6 | −0.575 | −0.309 | −0.333 | 0.097 | 0.171 | **0.629** | | | | | | | | | | | | |
| 7 | −0.584 | −0.366 | −0.356 | 0.147 | **0.231** | **0.255** | **0.685** | | | | | | | | | | | |
| 8 | **−8.252** | **−7.975** | **−6.735** | 0.371 | **1.134** | 0.300 | 0.387 | **5.392** | | | | | | | | | | |
| 9 | **−1.921** | **−1.520** | **−1.285** | 0.139 | **0.395** | **0.237** | **0.260** | **0.974** | **1.096** | | | | | | | | | |
| 10 | **−5.553** | **−5.628** | **−4.753** | 0.432 | **0.854** | 0.238 | 0.224 | **2.734** | **0.726** | **3.315** | | | | | | | | |
| 11 | −0.808 | −0.670 | −0.665 | 0.061 | 0.173 | 0.019 | 0.033 | 0.494 | **0.225** | 0.347 | **1.055** | | | | | | | |
| 12 | −0.207 | 0.651 | 0.895 | −0.009 | 0.022 | 0.106 | 0.121 | 0.134 | 0.167 | −0.056 | 0.115 | **1.420** | | | | | | |
| 13 | −0.811 | 0.540 | 1.080 | −0.225 | -0.080 | 0.225 | 0.173 | 0.418 | 0.251 | −0.021 | −0.008 | **1.070** | **2.210** | | | | | |
| 14 | **−1.394** | **−1.054** | −0.859 | −0.011 | 0.149 | 0.051 | −0.021 | **0.592** | **0.231** | **0.487** | 0.064 | 0.122 | 0.130 | **0.947** | | | | |
| 15 | −0.434 | 0.060 | 0.324 | −0.075 | 0.040 | 0.106 | 0.120 | 0.167 | 0.181 | 0.040 | 0.092 | **0.343** | **0.454** | 0.141 | **0.966** | | | |
| 16 | −0.289 | 0.035 | 0.161 | 0.011 | 0.044 | 0.083 | 0.067 | 0.129 | 0.179 | 0.083 | 0.123 | **0.298** | 0.336 | 0.119 | **0.525** | **0.976** | | |
| 17 | −0.472 | −0.042 | 0.120 | 0.019 | 0.057 | 0.109 | 0.118 | 0.255 | 0.173 | 0.084 | 0.106 | **0.321** | **0.427** | 0.074 | **0.526** | **0.538** | **0.937** | |
| 18 | **−2.789** | **−3.459** | **−3.278** | 0.144 | **0.287** | −0.045 | −0.054 | **1.220** | 0.120 | **0.923** | 0.046 | **−0.412** | **−0.520** | 0.142 | −0.216 | −0.206 | −0.192 | **1.079** |

*Note.* Estimates in bold have more than 95% of the posterior mass away from 0.

−6.269. This corresponds to an average price threshold of \$527. Because the maximum price of an APS camera is \$499, this implies that, on average, respondents are not screening on price. However, such analysis masks the tail behavior of the distribution of cutoff values. As displayed in Figure 2, 10% of the respondents will only consider cameras under \$276, 20% will only consider cameras under \$356, etc. In all, 41% of the respondents are screening alternatives based on price.

Figure 3 displays the proportion of respondents screening on each of the attributes. Body style, zoom, camera settings feedback, and price are the attributes most often used to screen alternatives. These attributes are available on both 35-mm and APS cameras, and as a result respondents are more fully aware of their benefits. In contrast, none of the new APS attributes (e.g., midroll film change) is being used to screen alternatives. In addition, the attributes used to form choice sets have substantively different part-worths in Tables 3 and 4. Our results indicate that once the choice set is formed, the price and body style do not play a role in the final decision. The zoom and camera feedback features, however, continue to play a role in the final selection of a camera beyond just screening. Previous authors have hypothesized that an attribute plays either a role in the screening phase or the choice phase (Bronnenberg and Vanhonacker 1996) or that once an attribute is deemed acceptable it will play no role in the final selection process (see Bettman 1979); these results provide evidence on this issue.

The participants in the survey were recruited and qualified by industry standards so that they were a prospect for cameras in the price range of the APS offering. Despite this prequalification, some 41% of respondents used price to screen alternatives and excluded those cameras above a particular threshold. Exploratory analysis revealed no relationship between the posterior estimate of the price threshold value and demographic variables, including income. However, replies to the question "Typically, when

**Table 4 Part-Worth Estimates for the Conjunctive Model**

| Attribute | Level | Beta | Posterior mean |
|---|---|---|---|
| Body style | Low body | 1 | −0.340 |
| | Medium body | 2 | 0.577 |
| | High body | 3 | 0.024 |
| Midroll change | Manual change | 4 | 0.257 |
| | Automatic change | 5 | **0.169** |
| Annotation | Preset list | 6 | **0.289** |
| | Customized list | 7 | **0.604** |
| | Input method 1 | 8 | −0.369 |
| | Input method 2 | 9 | **0.655** |
| | Input method 3 | 10 | **−0.619** |

| Attribute | Level | Beta | Posterior mean |
|---|---|---|---|
| Camera operation feedback | Operation feedback | 11 | **0.413** |
| Zoom | 2× zoom | 12 | **0.841** |
| | 4× zoom | 13 | **1.231** |
| Viewfinder | Large viewfinder | 14 | −0.139 |
| Settings feedback | LCD | 15 | **0.374** |
| | Viewfinder | 16 | **0.373** |
| | LCD & viewfinder | 17 | **0.520** |
| Price | ln(price) | 18 | 0.056 |

**Covariance Matrix—$\Sigma_\beta$**

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | **16.183** | | | | | | | | | | | | | | | | | |
| 2 | **15.035** | **15.918** | | | | | | | | | | | | | | | | |
| 3 | **14.839** | **15.176** | **16.419** | | | | | | | | | | | | | | | |
| 4 | −1.350 | −1.218 | −1.398 | **1.880** | | | | | | | | | | | | | | |
| 5 | **−1.903** | −1.821 | **−1.854** | **0.625** | **1.523** | | | | | | | | | | | | | |
| 6 | −0.637 | −0.527 | −0.591 | 0.117 | 0.148 | **0.632** | | | | | | | | | | | | |
| 7 | −0.667 | −0.531 | −0.604 | 0.198 | 0.194 | **0.241** | **0.742** | | | | | | | | | | | |
| 8 | **−6.744** | **−6.458** | **−6.301** | 0.653 | **0.956** | 0.349 | 0.399 | **4.828** | | | | | | | | | | |
| 9 | **−2.246** | **−2.069** | **−2.022** | 0.319 | **0.481** | **0.229** | **0.260** | **1.191** | **1.269** | | | | | | | | | |
| 10 | **−5.420** | **−5.306** | **−5.148** | 0.662 | **0.876** | 0.318 | 0.305 | **2.752** | **1.073** | **3.892** | | | | | | | | |
| 11 | −0.641 | −0.479 | −0.589 | 0.100 | 0.169 | 0.007 | 0.009 | 0.451 | 0.205 | 0.287 | **1.122** | | | | | | | |
| 12 | 0.443 | 0.499 | 0.491 | 0.035 | −0.051 | −0.005 | 0.006 | −0.119 | −0.048 | −0.161 | 0.126 | **1.056** | | | | | | |
| 13 | 0.106 | 0.427 | 0.387 | −0.101 | −0.179 | 0.112 | 0.080 | 0.084 | 0.014 | −0.234 | −0.015 | 0.557 | **1.653** | | | | | |
| 14 | **−1.292** | −1.256 | −1.289 | 0.057 | 0.154 | 0.065 | −0.047 | 0.616 | 0.281 | **0.695** | 0.022 | −0.049 | −0.112 | **1.213** | | | | |
| 15 | −0.648 | −0.611 | −0.494 | 0.005 | 0.070 | 0.045 | 0.079 | 0.280 | 0.125 | 0.186 | 0.087 | 0.097 | 0.139 | 0.102 | **0.864** | | | |
| 16 | −0.650 | −0.745 | −0.736 | 0.125 | 0.104 | 0.043 | 0.059 | 0.341 | 0.143 | 0.302 | 0.065 | 0.105 | 0.056 | 0.065 | **0.365** | **0.903** | | |
| 17 | −0.662 | −0.717 | −0.704 | 0.150 | 0.112 | 0.088 | 0.113 | 0.381 | 0.141 | 0.285 | 0.077 | 0.122 | 0.161 | 0.014 | **0.380** | **0.422** | **0.878** | |
| 18 | **−2.338** | **−2.463** | **−2.464** | 0.137 | 0.233 | 0.048 | 0.042 | **0.957** | 0.258 | **0.813** | 0.057 | −0.184 | −0.250 | 0.209 | 0.007 | 0.011 | 0.026 | **0.774** |

*Note.* Estimates in bold have more than 95% of the posterior mass away from 0.

you attend informal social gatherings away from your home, how often do you take a camera along with you?" were associated with the price threshold ($p$-value $= 0.0362$). Those respondents that indicated they "almost always" take their camera (44% of respondents) have an average price threshold \$68 higher than respondents who take their camera less frequently. Although exploratory in nature, this analysis shows that other variables correlate with the price threshold and may be used in conjunction with the model to better identify prospects in the marketplace.

The empirical results support the conjunctive screening rule model. Better in-sample and out-of-sample fit, despite the large increase in the number of parameters, argue for the use of the screening rule model over the more parsimonious probit. Consumers tend to screen alternatives using attributes they are already familiar with, and an exploratory analysis was conducted to demonstrate how the price threshold can be associated with other variables in a posterior analysis. After controlling for their effect in forming choice sets, the importance of the screening attributes in the final choice among acceptable alternatives is smaller than that indicated by a standard probit model.

## 7. Discussion

In this paper we have proposed a model to deal with an inherently discontinuous model of consumer behavior: two-stage decision making. Our reduced-form model incorporates screening rules that are consistent with various consideration set and information-processing theories of consumer behavior. We find support for the use of screening rules and for the conjunctive model in particular. The use of screening rules is pervasive, with 92% of respondents using this heuristic to manage the complexity of the choice problem. The empirical results also show that consumers screen alternatives using attributes that are well known, as opposed to the new and novel attributes also included in this study. Some attributes

**Table 5** **Threshold Estimates for the Conjunctive Model Posterior Means**

| Attribute | Levels | Recoded values | Possible cutoffs | Probability of each cutoff |
|---|---|---|---|---|
| Body style | Low | 0 | $\gamma_{1(1)} = -0.5$ | $\theta_1 = 0.598$ |
| | Medium | 1 | $\gamma_{1(2)} = 0.5$ | $\theta_2 = 0.324$ |
| | High | 2 | $\gamma_{1(3)} = 1.5$ | $\theta_3 = 0.059$ |
| | | | $\gamma_{1(4)} = 2.5$ | $\theta_4 = 0.015$ |
| Midroll change | None | 0 | $\gamma_{2(1)} = -0.5$ | $\theta_5 = 0.926$ |
| | Manual | 1 | $\gamma_{2(2)} = 0.5$ | $\theta_6 = 0.027$ |
| | Auto | 2 | $\gamma_{2(3)} = 1.5$ | $\theta_7 = 0.028$ |
| | | | $\gamma_{2(4)} = 2.5$ | $\theta_8 = 0.020$ |
| Annotation 1 | None | 0 | $\gamma_{3(1)} = -0.5$ | $\theta_9 = 0.940$ |
| | Preset | 1 | $\gamma_{3(2)} = 0.5$ | $\theta_{10} = 0.022$ |
| | Custom | 2 | $\gamma_{3(3)} = 1.5$ | $\theta_{11} = 0.018$ |
| | | | $\gamma_{3(4)} = 2.5$ | $\theta_{12} = 0.020$ |
| Annotation 2 | None | 0 | $\gamma_{4(1)} = -0.5$ | $\theta_{13} = 0.960$ |
| | Input 2 | 1 | $\gamma_{4(2)} = 0.5$ | $\theta_{14} = 0.020$ |
| | | | $\gamma_{4(3)} = 1.5$ | $\theta_{15} = 0.020$ |
| Annotation 3 | None | 0 | $\gamma_{5(1)} = -0.5$ | $\theta_{16} = 0.961$ |
| | Input 3 | 1 | $\gamma_{5(2)} = 0.5$ | $\theta_{17} = 0.019$ |
| | | | $\gamma_{5(3)} = 1.5$ | $\theta_{18} = 0.020$ |
| Continuous attribute: $-\ln(\text{price})$ | | | $\gamma_{11} = -6.269$ | |
| Annotation 4 | None | 0 | $\gamma_{6(1)} = -0.5$ | $\theta_{19} = 0.961$ |
| | Input #4 | 1 | $\gamma_{6(2)} = 0.5$ | $\theta_{20} = 0.019$ |
| | | | $\gamma_{6(3)} = 1.5$ | $\theta_{21} = 0.020$ |
| Operation feedback | None | 0 | $\gamma_{7(1)} = -0.5$ | $\theta_{22} = 0.947$ |
| | Feedback | 1 | $\gamma_{7(2)} = 0.5$ | $\theta_{23} = 0.034$ |
| | | | $\gamma_{7(3)} = 1.5$ | $\theta_{24} = 0.020$ |
| Zoom lens | None | 0 | $\gamma_{8(1)} = -0.5$ | $\theta_{25} = 0.808$ |
| | 2× | 1 | $\gamma_{8(2)} = 0.5$ | $\theta_{26} = 0.140$ |
| | 4× | 2 | $\gamma_{8(3)} = 1.5$ | $\theta_{27} = 0.032$ |
| | | | $\gamma_{8(4)} = 2.5$ | $\theta_{28} = 0.019$ |
| Viewfinder | Regular | 0 | $\gamma_{9(1)} = -0.5$ | $\theta_{29} = 0.961$ |
| | Large | 1 | $\gamma_{9(2)} = 0.5$ | $\theta_{30} = 0.019$ |
| | | | $\gamma_{9(3)} = 1.5$ | $\theta_{31} = 0.019$ |
| Settings feedback | None | 0 | $\gamma_{10(1)} = -0.5$ | $\theta_{32} = 0.783$ |
| | LCD | 1 | $\gamma_{10(2)} = 0.5$ | $\theta_{33} = 0.157$ |
| | Viewfinder | 2 | $\gamma_{10(3)} = 1.5$ | $\theta_{34} = 0.021$ |
| | Both | 3 | $\gamma_{10(4)} = 2.5$ | $\theta_{35} = 0.020$ |
| | | | $\gamma_{10(4)} = 3.5$ | $\theta_{36} = 0.019$ |

*Note.* All estimates have more than 95% of the posterior mass away from 0.

appear to be used only in forming choice sets, while others are used in both screening and final product choice. Although previous research has documented the improved fit of choice-set models and statistical biases that result from ignoring them, our method can apply a variety of screening rules and handle realistically sized problems. Our ability to capture a noncompensatory screening process and make attribute-level inferences for individuals is due to the unique model formulation and estimation technique.

A key element in our model is accurately navigating the likelihood of the data where only the final choice is observed. Past researchers have attempted to isolate the consideration-set probability Pr(C) in Equation (7) to create a smooth likelihood surface. This approach requires a complex partition of the error space corresponding to the $2^J - 1$ consideration sets and the choice probability within each set (see Mehta et al. 2003). Alternatively, the consideration-set probability can be estimated by treating each of the $2^J - 1$ partitions as a model parameter (see Chiang et al. 1999). By parameterizing the screening rules, using data augmentation to accurately partition the error space, and using the Markov chain to navigate the posterior distribution of all parameters, we can estimate a model that has a discontinuous likelihood function.

This research highlights the benefits and the potential of research on extended models of choice (Ben-Akiva et al. 1999). In this research we have demonstrated that the data are consistent with a two-stage decision process utilizing decision heuristics to form choice sets. In an exploratory analysis we show that the level of the price threshold was related to a camera usage question thought to encompass varying concerns and interests related to pursuing photography. It may very well be that the use of any decision heuristic, which decision heuristic (conjunctive or disjunctive), and/or which attributes to screen are also related to individual concerns and interests that may vary within an individual across usage occasions or choice contexts. For example, choosing a cam-

**Figure 2** **Distribution of Price Threshold**

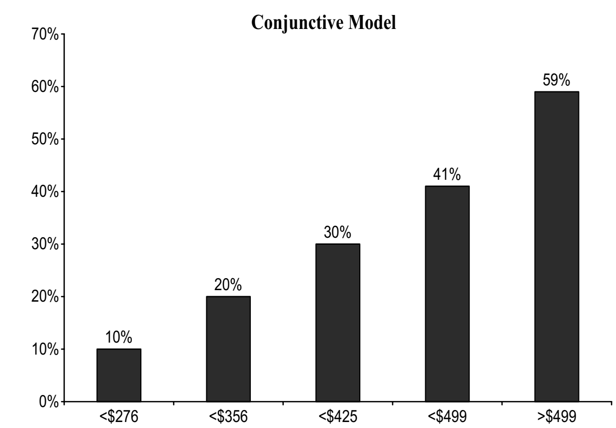
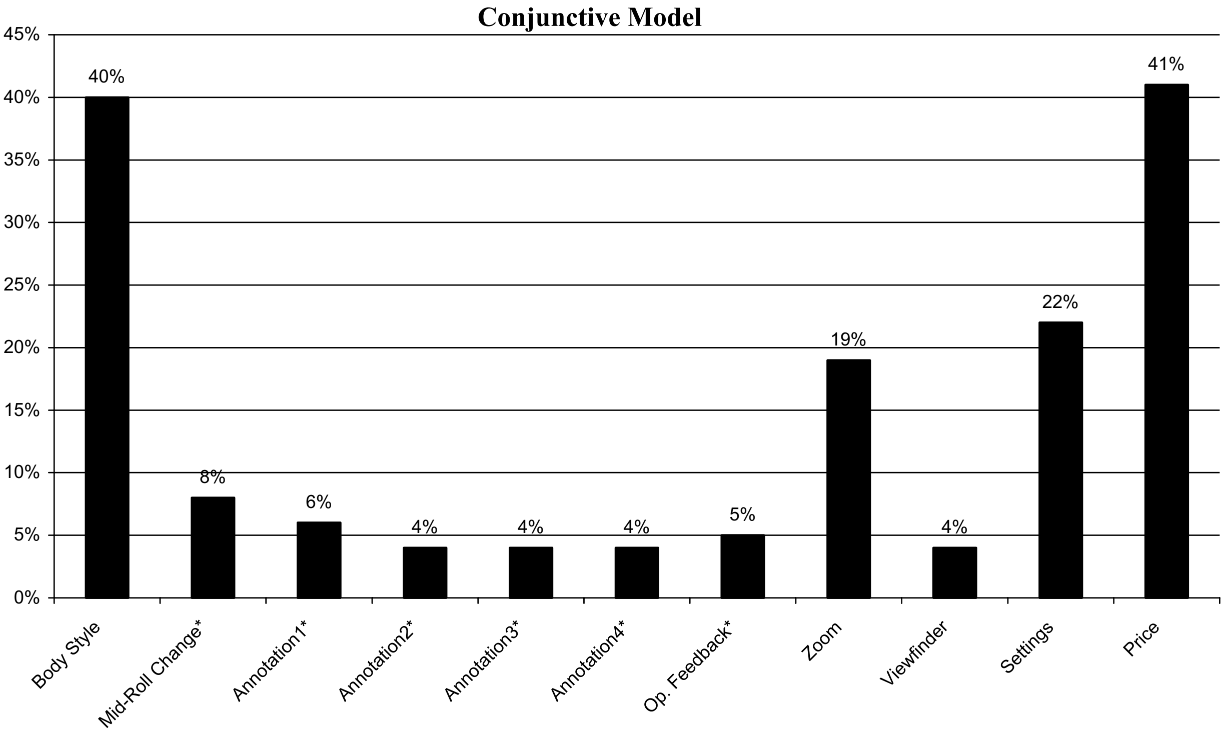


**Figure 3** **Proportion of Respondents Screening on Each Attribute**




*Note.* *Indicates attributes only available on APS cameras at the time of the study.

era as a gift for a loved one may result in more effortful decision making and less reliance on decision heuristics; choosing a single-use camera to take to the beach may be entirely characterized by a single-stage lexographic decision process (e.g., the cheapest). The method proposed in this paper can be used to test various discontinuous decision-making processes using revealed choice data.

The current study contains several limitations. As noted earlier, the compensatory model requires the deterministic component of utility to be the same in the screening rule and in the compensatory part of the model. More flexible representations that allow for a "screening utility" and a "choice utility" may fit the data better. Our empirical application may also be more prone to identifying screening rules than actual market situations. The large number of choice scenarios (14), the consistency of the universal set (six alternatives plus "none"), and the size of the universal set may all contribute to the adoption of screening rules by respondents. In these data, there was very little evidence that consumers used a disjunctive screening rule. It may be that more than 12 observations per individual are necessary to identify the parameters in this very flexible model.

The methodology presented in this paper can be extended. Attribute-level thresholds or the propensity to use one decision process over another in the structural heterogeneity model can be parameterized as functions of explanatory variables. The Bayesian hierarchical model would then be expanded by another level. The conjoint exercise used in this paper involved six alternatives described on eight attributes, plus a "none" option. Conjoint researchers may be interested to know if the number of alternatives or attributes influences respondents' use of a screening rule; this question can be studied using the proposed model and an experimental design that varies the number of attributes and alternatives. The model can also be applied to scanner panel datasets or any other panel dataset that captures repeated purchase behavior and identifies common attributes with unique levels. Such a study would reveal the importance of individual attributes and/or the use of screening rules in an actual market setting.

Managers frequently define markets based on particular attributes (e.g., body style) or price points, recognizing that certain consumers will only consider offerings with specific attribute levels, regardless of other elements of the product offering. Past methods of identifying these attributes and levels have tended to rely on compensatory models that may not be good representations of the decision process. While compensatory models can often predict as well as nonlinear models, inferences for strategic decisions are often biased. Our model overcomes the statistical and practical problems of estimating discontinuous choice models with attribute-level screening rules.

## Acknowledgments
The authors thank Mike Lotti and Eastman Kodak Company for providing the data used in this study.

## Appendix. Estimation Algorithm
Equations (13a)–(13c) list the conditional distributions for estimating the choice model in (12a)–(12b). The estimation for the conjunctive screening process, Equation (3), is described first. Extension to the disjunctive process, Equation (4), is straightforward. Changes necessary to estimate the compensatory process, Equations (2) and (13a′)–(13c′), are then described. The standard hierarchical Bayes probit model (Rossi et al. 1996) and structural heterogeneity models (Yang and Allenby 2000) have been described elsewhere. The augmented data problem for the conjunctive model is

$$y_{hij}=1 \quad \text{if } z_{hij}>z_{hik} \text{ for all } k \text{ such that } \mathrm{I}(x_{hik},\gamma_h)=1, \tag{A.1}$$

$$z_{hij}=x'_{hij}\beta_h+\varepsilon_{hij}, \quad \varepsilon\sim\text{Normal}(0,\sigma_\varepsilon=1), \tag{A.2}$$

$$\mathrm{I}(x_{hij},\gamma_h)=1 \quad \text{if } \prod_m \mathrm{I}(x_{hijm}>\gamma_{hm})=1, \tag{A.3}$$

where $h$ indexes the respondent, $i$ the choice set, $j$ the alternative in the choice set, and $m$ the attributes in each alternative. Heterogeneity on the individual household parameters is introduced hierarchically with the following distributions:

$$\beta_h \sim \text{MVN}(\bar\beta,\Sigma_\beta),$$

$$\gamma_{hm} \sim \text{Multinomial}(\theta_{m1},\ldots,\theta_{mn}) \quad \text{for discrete attributes } \gamma_1,\ldots,\gamma_{10},$$

$$\gamma_{h11} \sim \text{Normal}(\bar\gamma,\sigma^2_\gamma) \quad \text{for } \gamma_{11} \text{ a continuously scaled attribute.}$$

The hierarchy is completed by specifying the prior distributions on the hyperparameters:

$$\bar\beta \sim \text{MVN}(0,100\,\mathbb{1}) \qquad \Sigma_\beta\sim \text{IW}(\nu,\Delta),$$

$$\theta_m \sim \text{Dirichlet}(\alpha) \quad \text{for } m=1,\ldots,10, \text{ corresponding to discrete attributes},$$

$$\bar\gamma \sim \text{Normal}(0,100) \quad \sigma^2_\gamma\sim \text{IG}(a,b) \quad \text{for } \gamma_{11}, \text{ continuous attribute},$$

where IW is the inverted Wishart distribution, $\nu=k+8$, and $\Delta=\nu\,\mathbb{1}$. $\theta_m$ is a vector of dimension $n$ where $n$ indexes the possible values for $\gamma_m$ and $\alpha$ is a conforming vector with each element set equal to six. IG is the inverted gamma distribution and $a=10$ and $b=0.1$.

A Gibbs-type sampler is constructed to draw from the conditional posteriors described in (9a)–(9c) as well as the distributions for heterogeneity.

1. *Generate* $z_{hi}\mid y_{hi}, X_{hi}, \mathrm{I}(X_{hi},\gamma_h), \beta_h, \sigma_\varepsilon=1$ for $i=1,\ldots,I$ and $h=1,\ldots,H$.

Start with $j=1$,
if $y_{hij}=1$ then $z_{hij}\sim \text{TN}(x'_{hij}\beta_h, \sigma_\varepsilon=1, z_{hij}>z_{hjk}$ for all $k$ such that $\mathrm{I}(x_{hik},\gamma_h)=1)$,
if $y_{hij}=0$ and $\mathrm{I}(x_{hij},\gamma_h)=1$ then $z_{hij}\sim \text{TN}(x'_{hij}\beta_h, \sigma_\varepsilon=1, z_{hij}<z_{hik}$, where $y_{hik}=1)$,
else $z_{hij}\sim\text{Normal}(x'_{hij}\beta_h,\sigma_\varepsilon=1)$,
increment $j$ and return to top.

2. *Generate* $\gamma_{hm}\mid z_h, y_h, X_h$ for $m=1,\ldots,10$ (discrete attributes) and $h=1,\ldots,H$.

Conditioned on the current values for $z_h$ and the observed data, the set of allowable $\gamma^a_{hmn}$ must be identified from the set of possible $\gamma^p_{hmn}$. Here, $n$ indexes the possible values of $\gamma$ for each attribute as listed in Table 5 (e.g., $N=4$ for Body Style). Let $\mathrm{I}(\gamma^a_{hmn})=1$ indicate that $\gamma_{hmn}$ is a member of the set of allowable $\gamma_{hm}$.

Start with $n=1$, test each $j=1,\ldots,J$, $i=1,\ldots,I$,
if $y_{hij}=1$ and $\prod_m \mathrm{I}(x_{hijm}>\gamma^*_{hm})=0$, then $\mathrm{I}(\gamma^a_{hmn})=0$,
if $y_{hij}=0$ and $\prod_m \mathrm{I}(x_{hijm}>\gamma^*_{hm})=1$, then
if $z_{hij}>z_{hik}$ where $y_{hik}=1$, then $\mathrm{I}(\gamma^a_{hmn})=0$.
Else, $\mathrm{I}(\gamma^a_{hmn})=1$,
increment $n$ and return to the top.

Here $\gamma^*_{hm}$ includes $\gamma^p_{hmn}$ for the appropriate attribute. Now choose $\gamma_{hm}$ from $\gamma^a_{hmn}$ using a Metropolis-Hastings-type step:

$$\gamma_{hm}=\gamma^a_{hmn} \quad \text{with probability } \frac{\mathrm{I}(\gamma^a_{hmn})\theta_{mn}}{\sum_n \mathrm{I}(\gamma^a_{hmn})\theta_{mn}}.$$

This is an example of the "Griddy Gibbs" as described by Tanner (1993).

3. *Generate* $\gamma_{h11}\mid z_h, y_h, X_h$ for (continuous attribute) $h=1,\ldots,H$.

This is a more traditional random-walk Metropolis-Hastings step, but the candidate $\gamma^{(n)}_{h11}$, must be checked to make sure it is a member of the allowable set.

Let $\gamma^{(n)}_{h11}=\gamma^{(o)}_{h11}+e$ where $e\sim\text{Normal}(0,\sigma_e=0.5)$. The value $\sigma_e$ is chosen to ensure the acceptance rate is approximately 50%. $\gamma^{(o)}_{h11}$ is the value of $\gamma_{h11}$ from the previous iteration of the Gibbs sampler. Determine $\mathrm{I}(\gamma^{(n)a}_{h11})$ following the steps outlined in 2; if $\mathrm{I}(\gamma^{(n)a}_{h11})=0$, then $\gamma^{(n)}_{h11}=\gamma^{(o)}_{h11}$. If $\mathrm{I}(\gamma^{(n)a}_{h11})=1$, then accept $\gamma^{(n)}_{h11}$ with probability:

$$\min\left(\frac{\exp\left[-(2\sigma^2_\gamma)^{-1}\left(\bar\gamma-\gamma^{(n)}_{h11}\right)^2\right]}{\exp\left[-(2\sigma^2_\gamma)^{-1}\left(\bar\gamma-\gamma^{(o)}_{h11}\right)^2\right]},1\right).$$

4. *Generate* $\beta_h\mid X_h, z_h$ for $h=1,\ldots,H$.

$$\beta_h\sim\text{MVN}(b,(X'_hX_h+\Sigma^{-1}_\beta)^{-1}),$$

$$b=(X'_hX_h+\Sigma^{-1}_\beta)^{-1}(X'_hz_h+\Sigma^{-1}_\beta\bar\beta).$$

5. *Generate* $\bar\beta\mid\{\beta_h\},\Sigma_\beta$.

$$\bar\beta\sim\text{MVN}\left(\bar b,((\Sigma_\beta/H)^{-1}+(100\,\mathbb{1})^{-1})^{-1}\right),$$

$$\bar b=\left((\Sigma_\beta/H)^{-1}+(100\,\mathbb{1})^{-1}\right)^{-1}\left(\Sigma^{-1}_\beta\sum_{h=1}^H\beta_h+(100\,\mathbb{1})^{-1}(0)\right).$$

6. *Generate* $\Sigma_\beta\mid\{\beta_h\},\bar\beta$.

$$\Sigma_\beta\sim\text{IW}\left(\nu+H,\Delta+\sum_{h=1}^H(\beta_h-\bar\beta)'(\beta_h-\bar\beta)\right).$$

7. *Generate* $\theta_m\mid\{\gamma_{hm}\}$ for $m=1,\ldots,10$ (discrete attributes).

Define $s_{hmn}=1$ if $\gamma_{hm}=\gamma_{hmn}$ or $s_{hmn}=0$, otherwise. Then,

$$\theta_m\sim\text{Dirichlet}\left(\sum_{h=1}^H s_{hm1}+\alpha_1,\ldots,\sum_{h=1}^H s_{hmn}+\alpha_n\right).$$

See Allenby et al. (1998) for additional details on drawing $\theta_m$.

8. *Generate* $\bar{\gamma} \mid \{\gamma_{h11}\}, \sigma_\gamma^2$.

$$\bar{\gamma} \sim \text{Normal}(c, d),$$

$$c = \left( \frac{\sigma_\gamma^2/H}{\sigma_\gamma^2/H + 100}(0) + \frac{100}{\sigma_\gamma^2/H + 100} \left[ \frac{\sum_{h=1}^H \gamma_{h11}}{H} \right] \right)$$

$$d = \frac{100\sigma_\gamma^2/H}{\sigma_\gamma^2/H + 100}.$$

9. *Generate* $\sigma_\gamma^2 \mid \{\gamma_{h11}\}, \bar{\gamma}$.

$$\sigma_\gamma^2 \sim \text{IG}\left( \frac{H}{2} + a, \left\{ \frac{1}{b} + \frac{1}{2} \sum_{h=1}^H (\gamma_{h11} - \bar{\gamma})^2 \right\}^{-1} \right).$$

We now describe the changes necessary to estimate the compensatory consideration-set model. In our application of the model, the deterministic component of utility $V_{hij} = x'_{hij}\beta_h$ and (A.3) from above equal

$$\text{I}(x_{hij}, \gamma_h) = 1 \quad \text{if } \text{I}(x'_{hij}\beta_h > \gamma_h) = 1, \tag{A.3}$$

where $\gamma_h$ is a continuous scalar for each household with heterogeneity and prior distributions:

$$\gamma_h \sim \text{Normal}(\bar{\gamma}, \sigma_\gamma^2), \quad \bar{\gamma} \sim \text{Normal}(0, 100), \quad \sigma_\gamma^2 \sim \text{IG}(a, b).$$

Steps 3 and 7 from the previous procedure are no longer needed. Only Steps 2 and 4 need to be altered:

2a. *Generate* $\gamma_h \mid z_h, y_h, X_h, \beta_h$ for $h = 1, \ldots, H$.

Again, conditioned on the current values of $z_h$ and $\beta_h$ and the observed data, the key is to draw $\gamma_h$ from the set of allowable values. The allowable values of $\gamma_h$ define consideration sets such that the chosen alternative always has the maximum value of $z$.

If $y_{hij} = 1$, then $x'_{hij}\beta_h > \gamma_h$;
let $B^* = \min(x'_{hij}\beta_h$ for all $i$ and $j$, where $y_{hij} = 1)$.

If $y_{hij} = 0$ and $z_{hij} > z_{hik}$, where $y_{hik} = 1$, then $x'_{hij}\beta_h < \gamma_h$;
let $A^* = \max(x'_{hij}\beta_h$ for all $i$ and $j$ with $y_{hij} = 0$
and $z_{hij} > z_{hik}$, where $y_{hik} = 1)$.

Then $A^* < \gamma_h < B^*$ defines the allowable range for $\gamma_h$. $A^*$ requires that all chosen alternatives be part of the consideration set. $B^*$ requires that any alternative that is not chosen but has a value of $z$ greater than the chosen alternative, be excluded from the consideration set.

Following Chib and Greenberg (1995), we perform a Metropolis-Hastings step by drawing $\gamma_h$ directly from the allowable range of the prior; in our case the prior is the heterogeneity distribution and we draw from the truncated normal distribution:

$$\gamma_h \sim \text{TN}(\bar{\gamma}, \sigma_\gamma^2, A^* < \gamma_h < B^*).$$

4a. *Generate* $\beta_h \mid y_h, X_h, z_h$ for $h = 1, \ldots, H$.

Generate a candidate $\beta_h$:

$$\beta_h^c \sim \text{MVN}(b, (X'_h X_h + \Sigma_\beta^{-1})^{-1}),$$

$$b = (X'_h X_h + \Sigma_\beta^{-1})^{-1}(X'_h z_h + \Sigma_\beta^{-1}\bar{\beta}).$$

Define $A^*$ and $B^*$ as in Step 2a with $\beta_h^c$. If $A^* > B^*$ reject $\beta_h^c$, draw a new candidate. If $A^* < B^*$, then $\beta_h = \beta_h^c$. This is an example of acceptance/rejection sampling.

## References

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

Allenby, Greg M., Peter E. Rossi. 1999. Marketing models of consumer heterogeneity. *J. Econometrics* **89** 57–78.

Allenby, Greg M., Neeraj Arora, James L. Ginter. 1998. On the heterogeneity of demand. *J. Marketing Res.* **35**(August) 384–389.

Allenby, Greg, Neeraj Arora, Chris Diener, Jaehwan Kim, Mike Lotti, Paul Markowitz. 2002. Distinguishing likelihoods, loss functions, and heterogeneity in the evaluation of marketing models. *Canadian J. Marketing Res.* **20** 44–59.

Andrews, Rick L., Ajay K. Manrai. 1998. Simulation experiments in choice simplification: The effects of task and context on forecasting performance. *J. Marketing Res.* **35**(May) 198–209.

Andrews, Rick L., T. C. Srinivasan. 1995. Studying consideration effects in empirical choice models using scanner panel data. *J. Marketing Res.* **32**(February) 30–41.

Ben-Akiva, Moshe, Bruno Boccara. 1995. Discrete choice models with latent choice sets. *Internat. J. Res. Marketing* **12** 9–24.

Ben-Akiva, Moshe, Daniel McFadden, Tommy Garling, Dinesh Gopinath, Joan Walker, Denis Bolduc, Axel Borsch-Supan, Philippe Delquie, Oleg Larichev, Taka Morikawa, Amalia Polydoropoulou, Vithala Rao. 1999. Extended framework for modeling choice behavior. *Marketing Lett.* **10** 187–203.

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

Bettman, James R., C. Whan Park. 1980. Effects of prior knowledge and experience and phase of the choice process on consumer decision analysis: A protocol analysis. *J. Consumer Behavior* **7** 234–248.

Bettman, James R., Eric J. Johnson, John W. Payne. 1990. A componential analysis of cognitive effort in choice. *Organ. Behavior Human Decision Processes* **45** 111–139.

Bettman, James R., Mary Frances Luce, John W. Payne. 1998. Constructive consumer choice processes. *J. Consumer Res.* **25** 187–217.

Bronnenberg, Bart J., Wilfried R. Vanhonacker. 1996. Limited choice sets, local price response, and implied measures of price competition. *J. Marketing Res.* **33**(May) 163–173.

Chiang, Jeongwen, Siddhartha Chib, Chakravarthi Narasimhan. 1999. Markov chain Monte Carlo and models of consideration set and parameter heterogeneity. *J. Econometrics* **89** 223–248.

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

Edwards, Yancy D., Greg M. Allenby. Multivariate analysis of multiple response data. *J. Marketing Res.* **40**(August) 321–334.

Fader, Peter S., Leigh McAlister. 1990. An elimination by aspects model of consumer response to promotion calibrated on UPC scanner data. *J. Marketing Res.* **27**(August) 322–332.

Gensch, Dennis H. 1987. A two-stage disaggregate attribute choice model. *Marketing Sci.* **6**(3) 223–239.

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

Häubl, Gerald, Valerie Trifts. 2000. Consumer decision making in online shopping environments: The effects of interactive decision aids. *Marketing Sci.* **19**(1) 4–21.

Hauser, John R., Birger Wernerfelt. 1990. An evaluation cost model of consideration sets. *J. Consumer Res.* **16**(March) 393–408.

Jedidi, Kamel, Rajeev Kohli, Wayne S. DeSarbo. 1996. Consideration sets in conjoint analysis. *J. Marketing Res.* **33**(August) 364–372.

Johnson, Eric J., Robert J. Meyer. 1984. Compensatory choice models of noncompensatory processes: The effect of varying context. *J. Consumer Res.* **11** 528–541.

Johnson, Eric J., Robert J. Meyer, Sanjoy Ghose. 1989. When choice models fail: Compensatory models in negatively correlated environments. *J. Marketing Res.* **26**(August) 255–270.

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

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

Manrai, Ajay K., Rick L. Andrews. 1998. Two-stage choice models for scanner panel data: An assessment of process and assumptions. *Eur. J. Oper. Res.* **111** 193–215.

McCulloch, Robert, Peter E. Rossi. 1994. An exact likelihood analysis of the multinomial probit model. *J. Econometrics* **64** 207–240.

Mehta, Nitin, Surendra Rajiv, Kannan Srinivasan. 2003. Price uncertainty and consumer search: A structural model of consideration set formation. *Marketing Sci.* **22**(1) 58–84.

Montgomery, Alan J., Katrik Hosanagar, Ramayya Krishnan, Karen B. Clay. 2003. Designing a better shopbot. Working paper, Carnegie Mellon University, Pittsburgh, PA.

Newton, Michael A., Adrian E. Raftery. 1994. Approximating Bayesian inference with the weighted likelihood bootstrap. *J. Roy. Statist. Soc. B* **56** 3–48.

Payne, John W. 1976. Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organ. Behavior Human Performance* **16** 366–387.

Roberts, John H., James M. Lattin. 1991. Development and testing of a model of consideration set compostion. *J. Marketing Res.* **28**(November) 429–440.

Rossi, Peter E., Robert E. McCulloch, Greg M. Allenby. 1996. The value of purchase history data in target marketing. *Marketing Sci.* **15**(4) 321–340.

Shocker, Allan D., Moshe Ben-Akiva, Bruno Boccara, Prakash Nedungandi. 1991. Consideration set influences on consumer decision making and choice: Issues, models, and suggestions. *Marketing Lett.* **2**(3) 181–197.

Shugan, Steven M. 1980. The cost of thinking. *J. Consumer Res.* **7**(September) 99–111.

Siddarth, S., Randolph E. Bucklin, Donald G. Morrison. 1995. Making the cut: Modeling and analyzing choice set restriction in scanner panel data. *J. Marketing Res.* **32**(August) 255–266.

Simmons Study of Media and Markets. 1997. Choices II Version 5.28.98. Simmons Market Research Bureau, Tampa, FL.

Swait, Joffre. 2001. A non-compensatory model incorporating attribute cutoffs. *Transportation Res. Part B* **35** 903–928.

Tanner, Martin A. 1993. *Tools for Statistical Inference*. Springer-Verlag, New York.

Tanner, M. A., W. Wong. 1987. The calculation of posterior distributions by data augmentation. *J. Amer. Statist. Assoc.* **82** 528–549.

Yang, Sha, Greg M. Allenby. 2000. A model for observation, structural, and household heterogeneity in panel data. *Marketing Lett.* **11** 137–149.