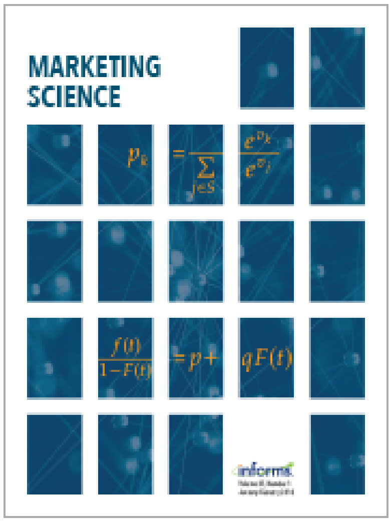



Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

# Efficient Choice Designs for a Consider-Then-Choose Model

Qing Liu, Neeraj Arora,

To cite this article:
Qing Liu, Neeraj Arora, (2011) Efficient Choice Designs for a Consider-Then-Choose Model. Marketing Science 30(2):321-338.
https://doi.org/10.1287/mksc.1100.0629

MARKETING SCIENCE
CELEBRATING 30 YEARS
Vol. 30, No. 2, March–April 2011, pp. 321–338
ISSN 0732-2399 | EISSN 1526-548X | 11 | 3002 | 0321

informs®
DOI 10.1287/mksc.1100.0629


# Efficient Choice Designs for a Consider-Then-Choose Model

Qing Liu, Neeraj Arora
Department of Marketing, University of Wisconsin–Madison, Madison, Wisconsin 53706
{qliu@bus.wisc.edu, narora@bus.wisc.edu}

Existing research on choice designs focuses exclusively on compensatory models that assume that all available alternatives are considered in the choice process. In this paper, we develop a method to construct efficient designs for a two-stage, consider-then-choose model that involves a noncompensatory screening process at the first stage and a compensatory choice process at the second stage. The method applies to both conjunctive and disjunctive screening rules. Under certain conditions, the method also applies to the subset conjunctive and disjunctions of conjunctions screening rules. Based on the local design criterion, we conduct a comparative study of compensatory and conjunctive designs—the former are optimized for a compensatory model and the latter for a two-stage model that uses conjunctive screening in its first stage. We find that conjunctive designs have higher level overlap than compensatory designs. This occurs because level overlap helps pinpoint screening behavior. Higher overlap of conjunctive designs is also accompanied by lower orthogonality, less level balance, and more utility balance. We find that compensatory designs have a significant loss of design efficiency when the true model involves conjunctive screening at the consideration stage. These designs also have much less power than conjunctive designs in identifying a true consider-then-choose process with conjunctive screening. In contrast, when the true model is compensatory, the efficiency loss from using a conjunctive design is lower. Also, conjunctive designs have about the same power as compensatory designs in identifying a true compensatory choice process. Our findings make a strong case for the use of conjunctive designs when there is prior evidence to support respondent screening.

*Key words*: experimental design; conjoint choice designs; *D*-optimality; consider-then-choose model; noncompensatory screening rules
*History*: Received: July 22, 2009; accepted: November 27, 2010; Eric Bradlow served as the editor-in-chief and Michel Wedel served as associate editor for this article. Published online in *Articles in Advance* February 9, 2011.

## 1. Introduction

Choice-based conjoint experiments are widely used in marketing to measure consumer preference. A majority of these experiments rely on a compensatory utility specification. Recent innovations in preference measurement point to limitations of the compensatory utility specification and therefore advocate models based on a two-stage, consider-then-choose decision process (e.g., Gilbride and Allenby 2004, Kohli and Jedidi 2005, Hauser et al. 2010). The basic premise of these models is that consumers use noncompensatory decision rules to screen out certain alternatives from consideration before engaging in a compensatory trade-off process. A significant hurdle in the adoption of such consider-then-choose models is that there is no guidance on how to design a choice experiment for such models. This is in contrast with compensatory choice models, for which it is well known how efficient designs could be constructed (e.g., Sándor and Wedel 2002, Kessels et al. 2006, Yu et al. 2009). Our goal in this paper is to fill this void by providing a statistical methodology to help design experiments for a consider-then-choose model that involves a noncompensatory screening process at the consideration stage. Using this methodology, we investigate questions of design size, characteristics, efficiency, and robustness.

Extant research on consideration set formation provides strong evidence that consumers use simplifying heuristics before making a choice. For example, after reviewing a number of studies, Bettman (1979) concludes that consumers often use a conjunctive screening rule to reduce the number of products before engaging in a compensatory evaluation of the final choice set under consideration. In the conjunctive screening rule, a product is included for consideration if *all* features (e.g., color, brand, and price level) of the product are acceptable. Disjunctive screening is an alternative rule whereby a product is considered if at least *one* feature (e.g., color) of the product is acceptable. Other known noncompensatory screening rules include subset conjunctive (Kohli and

Jedidi 2005) and disjunctions of conjunctions (Hauser et al. 2010). Both these screening rules involve a combination of conjunctive and disjunctive rules on subsets of product features. Specifically, in the subset conjunctive rule with subset size $k$, a product is included for consideration if at least $k$ features of the product are acceptable. In the disjunctions of conjunctions (DOC) screening rule, a product is considered if one or more "patterns," or conjunctions of screening criteria, are satisfied.

For the class of consider-then-choose models that are the focus of this paper, the conjunctive screening rule has generated significant attention in the literature. For example, Hauser et al. (2010, p. 487) find in their in-depth interviews that "typically, consumers used conjunctive-like criteria." Gilbride and Allenby (2004) find that the conjunctive screening model outperforms the disjunctive model and a mixture of conjunctive and disjunctive models. Furthermore, evidence for the use of the conjunctive screening rule in consumer choice process has been documented in a variety of fields that include health-care economics, urban planning, and marketing (e.g., Payne 1976, Bettman 1979, Lussier and Olshavsky 1979, Swait 2001, Araña et al. 2008, Zhu and Timmermans 2009).

In this paper, we develop a method to construct efficient designs for a two-stage, consider-then-choose model that involves a noncompensatory screening process at the consideration stage. It applies to conjunctive and disjunctive screening rules. It can also be applied to the subset conjunctive screening rule if the subset size $k$ is known a priori and to the disjunctions of conjunctions screening rule if the patterns are known a priori. We use the modeling framework in Gilbride and Allenby (2004, 2006) to account for these various screening rules. Using the Fisher information matrix, we derive the general form of the design criterion to construct designs optimized for the consider-then-choose model. We show that this design criterion is very different from the one for a compensatory choice model.

In light of the emphasis placed on the conjunctive screening rule by existing literature, in this paper, we focus our attention on the consider-then-choose model with the conjunctive screening rule at the consideration stage.[1] We refer to the designs optimized for the conjunctive model as conjunctive designs. We conduct a comprehensive comparative study that contrasts conjunctive designs with compensatory designs on a variety of key issues that include design size, characteristics, efficiency, and robustness. In particular, we address the following questions as they relate to each issue.

(1) *Design size*: Are conjunctive designs significantly larger than their compensatory counterparts? In particular, for a fixed number of alternatives per choice set, what is the minimum number of choice sets that should be included in the choice task? Do conjunctive designs require more choice sets than compensatory designs to achieve the same amount of information?

(2) *Design characteristics*: Extant research on compensatory choice designs has emphasized the following four design aspects: level overlap, orthogonality, level balance, and utility balance. How do conjunctive designs differ from compensatory designs on these four aspect?

(3) *Design efficiency*: How efficient are the compensatory designs when the true choice decision process is conjunctive? Conversely, how efficient are conjunctive designs when the true choice decision process is compensatory? Which type of design has more power in identifying the true underlying compensatory or conjunctive process?

(4) *Design robustness*: How sensitive are conjunctive designs to the prior assumptions about the model parameters? In particular, how would the design efficiency change if the assumed parameter values used to construct the conjunctive design deviate from the true values?

Results from our comparative study reveal several interesting insights that make a strong case for using conjunctive designs. First, we find that, ceteris paribus, the minimum number of choice sets required for a conjunctive design is not any higher than that for a compensatory design when the extent of screening is low. However, when a higher proportion of individuals engage in screening across multiple product features, the minimum number of choice sets required for a conjunctive design is larger. Second, the characteristics of conjunctive designs are quite different from compensatory designs—conjunctive designs exhibit higher level overlap. This occurs because level overlap helps pinpoint screening behavior, and designs with higher overlap are therefore more efficient when the choice process involves a screening stage. Higher overlap also results in lower orthogonality, less level balance, and more utility balance of conjunctive designs. Third, compensatory designs have significant loss of design efficiency when the true model is conjunctive. In contrast, when the true model is compensatory, the efficiency loss from using a conjunctive design is significantly lower. Fourth, although compensatory designs have much less power than conjunctive designs do in identifying a true conjunctive process, surprisingly, conjunctive designs are found to have about the same power as compensatory designs in identifying a true compensatory choice process. Finally, robustness checks on

[1] For simplicity, we refer to this model as the conjunctive model in the rest of this paper.

conjunctive designs indicate that they are fairly robust to the assumptions of the covariance matrix of the random effects. However, efficiency decreases when the assumed preference and screening parameter values deviate from the true values.

The rest of this paper is organized as follows. In §2 we review the existing compensatory design criteria for a hierarchical choice model with uncorrelated and correlated random effects. In §3 we present a two-stage, consider-then-choose model that accounts for a variety of screening rules and derive the design criterion for the model. In §4 we describe the algorithm used in the design construction and the scenarios we use for our comparative study that contrasts conjunctive designs with compensatory designs. Results of our comparative study are reported in §5. In §6 we discuss the sensitivity of conjunctive designs to prior specifications of model parameters. We conclude with a summary in §7 and a discussion for future research in §8.

## 2. Compensatory Design Criteria

Existing research on choice designs assumes a compensatory model structure. Based on McFadden's (1974) random utility model of consumer choice, the utility for an alternative $j$ in choice set $s$ for a given person is

$$u_{sj} = x'_{sj}\beta + e_{sj}, \tag{1}$$

where $x_{sj}$ is a vector denoting the features of the alternative, $\beta$ is a parameter vector reflecting the person's preferences, and the errors $\{e_{sj}\}$ are assumed to have independent Gumbel (or type 1 extreme value) distribution with location parameter 0 and scale parameter 1. The no-choice option, when included, is treated as a separate alternative with zero utility by setting the vector $x_{sj} = 0$. For commonly used hierarchical models that account for consumer heterogeneity (e.g., Allenby et al. 1998, Wedel et al. 1999, Bradlow and Rao 2000), we begin with the design criteria for a compensatory, mixed-logit choice model that uses a random effects specification of $\beta$.

### 2.1. Existing Criteria for the Special Case of Uncorrelated Random Effects

Building upon earlier research on choice designs (Huber and Zwerina 1996, Arora and Huber 2001, Kuhfeld and Tobias 2005), Sándor and Wedel (2002) proposed the use of the local $D$-criterion for the construction of efficient designs for hyperparameter estimation under a mixed-logit model with uncorrelated random effects. The mixed-logit model assumes that the preference parameters in vector $\beta$ are random effects from a multivariate normal distribution with mean $\mu_B$ and a diagonal covariance matrix $\Lambda = \text{diag}(\sigma_1^2, \ldots, \sigma_m^2)$. That is, $\beta = \mu_B + V\sigma_B$ with vector $\sigma_B = (\sigma_1, \ldots, \sigma_m)'$ , and $V$ is an $m \times m$ diagonal matrix whose $m$ diagonal elements are independent and identically distributed from the standard normal distribution. If we let vector $\nu$ represent these $m$ standard normal variables, the probability that alternative $j$ is chosen from choice set $s$, given $\mu_B$ and $\sigma_B$, is

$$\varphi_{sj} = \int p_{sj}(\nu) f(\nu)\, d\nu, \quad \text{where}$$

$$p_{sj}(\nu) = \frac{\exp\{x'_{sj}(\mu_B + V\sigma_B)\}}{\sum_{j=1}^{J} \exp\{x'_{sj}(\mu_B + V\sigma_B)\}}. \tag{2}$$

Define $p_s(\nu) = (p_{s1}(\nu), \ldots, p_{sJ}(\nu))'$, $P_s(\nu) = \text{diag}(p_{s1}(\nu), \ldots, p_{sJ}(\nu))$, $\Delta_s = \text{diag}(\varphi_{s1}, \ldots, \varphi_{sJ})$, and let $X_s = (x_{s1}, x_{s2}, \ldots, x_{sJ})'$ represent the design matrix for choice set $s$. The local $D$-optimal designs for efficient estimation of hyperparameter vectors $\mu_B$ and $\sigma_B$ are constructed by minimizing the determinant of the inverse of the Fisher information matrix $I(\mu_B, \sigma_B)$ normalized by the total number of the hyperparameters (which is $2m$ in this case). That is, the local $D$-criterion is defined as the minimization of

$$D\text{-error} = \det\{I(\mu_B, \sigma_B)^{-1}\}^{1/2m}, \tag{3}$$

where

$$I(\mu_B, \sigma_B) = N \sum_{s=1}^{S} \begin{bmatrix} E'_s \Delta_s^{-1} E_s & E'_s \Delta_s^{-1} Q_s \\ Q'_s \Delta_s^{-1} E_s & Q'_s \Delta_s^{-1} Q_s \end{bmatrix}, \tag{4}$$

with

$$E_s = \int [P_s(\nu) - p_s(\nu)p_s(\nu)'] X_s f(\nu)\, d\nu,$$

$$Q_s = \int [P_s(\nu) - p_s(\nu)p_s(\nu)'] X_s V f(\nu)\, d\nu.$$

Yu et al. (2009) extended the local $D$-criterion to the corresponding full Bayesian and semi-Bayesian design criteria by taking into account the uncertainty of the hyperparameters through the prior probability density functions $f(\mu_B, \sigma_B)$ (for the full Bayesian criterion) and $f(\mu_B)$ (for the semi-Bayesian criterion). However, the search for optimal semi-Bayesian and full Bayesian designs is computationally demanding, and Yu et al. (2009) rely on supercomputers to obtain such designs. Given the computational constraints that currently exist, we focus on local designs in this paper.

### 2.2. Extension to the General Case of Correlated Random Effects

To set a foundation for optimal designs for a consider-then-choose model, next we extend the local compensatory design criterion to the more general scenario when the random effects may be correlated. That is, the covariance matrix $\Lambda$ is of a general form where the off-diagonal elements may not be zero. We let $\lambda_B$ denote the vector of the $m(m+1)/2$ unique parameters in $\Lambda$. We also allow the inclusion of the covariates $Z_i$ such that the random effects associated with respondent $i$ are multivariate normal with mean $Z_i\theta$ and covariance matrix $\Lambda$. The covariate matrix $Z_i$ is of size $m \times q$, and the hyperparameter vector $\theta$ is of length $q$. Let vector $\varepsilon \sim$ Multivariate Normal$(0, \Lambda)$. The probability that alternative $j$ is chosen by respondent $i$ from choice set $s$, given $\theta$ and $\lambda_B$, is therefore

$$\varphi_{isj} = \int p_{isj}(\varepsilon) f(\varepsilon \mid \Lambda)\, d\varepsilon, \quad \text{where}$$
$$p_{isj}(\varepsilon) = \frac{\exp\{x'_{sj}(Z_i\theta + \varepsilon)\}}{\sum_{j=1}^{J} \exp\{x'_{sj}(Z_i\theta + \varepsilon)\}}. \tag{5}$$

The Fisher information matrix for hyperparameter vectors $\theta$ and $\lambda_B$ can be expressed as (see Appendix A for details)

$$I(\theta, \lambda_B) = \sum_{i=1}^{N}\sum_{s=1}^{S} \begin{bmatrix} A'_{is}\Delta_{is}^{-1}A_{is} & A'_{is}\Delta_{is}^{-1}\Gamma_{is} \\ \Gamma'_{is}\Delta_{is}^{-1}A_{is} & \Gamma'_{is}\Delta_{is}^{-1}\Gamma_{is} \end{bmatrix}. \tag{6}$$

Under this general setup of the compensatory hierarchical choice model, the total number of hyperparameters is $c = q$ (for $\theta$) $+ m(m+1)/2$ (for $\Lambda$). The local $D$-criterion is therefore the minimization of

$$D^{\text{comp}}\text{-error} = \det\{I(\theta, \lambda_B)^{-1}\}^{1/c}, \quad \text{where}$$
$$c = q + m(m+1)/2. \tag{7}$$

## 3. Consider-Then-Choose Model and Design Criterion

In this section we introduce the local design criterion for a consider-then-choose model that accounts for a variety of screening rules. Compared with its compensatory counterpart, the local design criterion for the consider-then-choose model involves additional screening parameters as outlined in §3.1.

### 3.1. Consider-Then-Choose Model with Noncompensatory Screening

The consider-then-choose model is fairly flexible and can account for different forms of screening for nominal attributes commonly used in marketing applications. We use the same framework as in Gilbride and Allenby (2004, 2006). For a given respondent, let $\tau_r$ denote whether an attribute level $r$ is acceptable ($\tau_r = 0$) or not ($\tau_r = 1$). Given the screening indicators $\boldsymbol{\tau} = \{\tau_r\}$ on all attribute levels, an alternative is included for consideration if and only if it satisfies the screening rule used in the formation of the consideration set. Specifically, alternative $j$ in choice set $s$ is included for consideration if and only if

$$\begin{cases} \sum_{r \in AL(j;\,s)} I\{\tau_r = 0\} \geq 1, \\ \quad \text{for disjunctive screening,} \\ \sum_{r \in AL(j;\,s)} I\{\tau_r = 0\} = R_j, \ \text{or} \ \prod_{r \in AL(j;s)} I\{\tau_r = 0\} = 1, \\ \quad \text{for conjunctive screening,} \\ \sum_{r \in AL(j;\,s)} I\{\tau_r = 0\} \geq k,\ k < R_j, \\ \quad \text{for subset conjunction,} \\ \sum_{par=1}^{P} \left( \prod_{r \in L_{par}(j;\,s)} I\{\tau_r = 0\} \right) \geq 1, \\ \quad \text{for disjunctions of conjunctions,} \end{cases} \tag{8}$$

where $I\{\cdot\}$ is the indicator function, $AL(j; s)$ consists of the attribute levels of alternative $j$ in choice set $s$, $R_j$ is the total number of attribute levels present in alternative $j$, and $L_{par}(j; s)$ consists of the attribute levels of alternative $j$ in choice set $s$ that are present in pattern $par$. The expressions in (8) capture four different screening rules. In particular, an alternative is included in the consideration set if

(i) At least *one* feature is acceptable (for disjunctive screening); for example, a flat panel TV is considered acceptable if the brand is Sony or the screen type is LCD or the resolution is 1,080p or the price is less than \$1,500.

(ii) *All* features are acceptable (for conjunctive screening); for example, a flat panel TV is considered acceptable if the brand is Sony and the screen type is LCD and the resolution is 1,080p and the price is less than \$1,500.

(iii) At least $k$ features are acceptable (for subset conjunctive); for example, a flat panel TV is considered acceptable if at least two ($k = 2$) criteria (e.g., any two from the following list: Sony brand, LCD screen, 1,080p resolution, or price less than \$1,500) are satisfied.

(iv) One or more *patterns* or conjunctions are satisfied (for disjunctions of conjunctions); for example, one pattern may be "Sony brand and LCD screen and price less than \$1,500," and another pattern may be "LCD screen and 1,080p resolution."

For a given choice set, if all alternatives in the choice set are screened out, then the consideration set is empty and the no-choice option is selected with probability 1. If the consideration set is not empty, then the alternatives included in the consideration set, together with the no-choice option, are selected with probabilities according to the multinomial logit model. All remaining alternatives are selected with probability 0. In particular, let $C_{is}(\tau)$ denote the consideration set for respondent $i$ and choice set $s$, where

$C_{is}(\tau) = \varnothing$ denotes the case when the consideration set is empty. Let $j = 0$ represent the no-choice option; then vector $x_{sj} = 0$ when $j = 0$. Given the consideration set $C_{is}(\tau)$, the conditional probability that alternative $j$ is chosen by respondent $i$ from choice set $s$ can be expressed as

$$\begin{aligned}
&P(y_{is} = 0 \mid C_{is}(\tau) = \varnothing) = 1,\\
&P(y_{is} = j \mid j \neq 0, C_{is}(\tau) = \varnothing) = 0,\\
&P(y_{is} = j \mid j \neq 0, j \notin C_{is}(\tau), C_{is}(\tau) \neq \varnothing) = 0,\\
&P(y_{is} = j \mid j = 0 \text{ or } j \in C_{is}(\tau), C_{is}(\tau) \neq \varnothing, \theta, \varepsilon)\\
&\quad = \frac{\exp\{x'_{sj}(Z_i\theta + \varepsilon)\}}{\sum_{j=0 \text{ or } j\in C_{is}(\tau)} \exp\{x'_{sj}(Z_i\theta + \varepsilon)\}},
\end{aligned} \tag{9}$$

where $\varepsilon \sim$ Multivariate Normal$(0, \Lambda)$, and $Z_i\theta + \varepsilon$ denotes the random preference effects that follow a multivariate Normal distribution with mean $Z_i\theta$ and covariance matrix $\Lambda$. The screening indicators are also assumed to be random effects such that $\tau_r \sim$ Bernoulli$(\xi_r)$. Although a Bernoulli specification is used in this paper, we note that the same methodology for obtaining the Fisher information matrix (see Appendix B for details) applies to a multinomial specification for nominal attributes (Gilbride and Allenby 2004). Choice decisions are independent across choice sets and respondents.

### 3.2. Design Criterion for the Consider-Then-Choose Model

For the consider-then-choose model with noncompensatory screening described in §3.1, we next obtain the criterion for constructing choice designs that are efficient for the estimation of the hyperparameters $\theta$, $\Lambda$, and $\zeta$, where $\zeta = \{\xi_r\}$. It is evident that the consider-then-choose design criterion involves an additional $R$ screening parameters in vector $\zeta$. The presence of these screening parameters complicates the Fisher information matrix considerably. For example, if the choice task consists of four attributes each with three levels, then the total number of attribute levels equals $R = 4 \times 3 = 12$, and correspondingly there are 12 screening parameters $\zeta = \{\xi_1, \ldots \xi_{12}\}$. The screening indicator $\tau_r$ for each of the attribute levels $r(r = 1, \ldots, R)$ can take on one of two values (0 or 1) and thus results in a total of $2^R$ possible screening scenarios. The first scenario is when all indicators $\tau_r$ equal 0; that is, when all attribute levels are accepted. The last scenario is when all indicators $\tau_r$ equal 1; that is, when all attribute levels are screened out. We denote these scenarios as $\boldsymbol{\tau}^1, \boldsymbol{\tau}^2, \ldots, \boldsymbol{\tau}^{2^R}$. The probability that alternative $j$ is chosen by respondent $i$ from choice set $s$, given the hyperparameters $\theta$, $\lambda_B$, and $\zeta$, is therefore

$$\begin{aligned}
\pi_{isj} = \int \sum_{\tau=\tau^1}^{\tau^{2^R}} &\left\{ p_{isj}(\tau, \varepsilon) \prod_{r=1}^{R} \xi_r^{I\{\tau_r=1\}} (1-\xi_r)^{I\{\tau_r=0\}} \right\}\\
&\cdot f(\varepsilon \mid \Lambda)\, d\varepsilon,
\end{aligned} \tag{10}$$

where $p_{isj}(\tau, \varepsilon)$ is the conditional probability defined according to Expression (9). The Fisher information matrix $I(\theta, \lambda_B, \zeta)$ for the consider-then-choose model has the following expression (see Appendix B for details):

$$\begin{aligned}
&I(\theta, \lambda_B, \zeta)\\
&= \sum_{i=1}^{N}\sum_{s=1}^{S} \begin{pmatrix} M'_{is}\Pi_{is}^{-1}M_{is} & M'_{is}\Pi_{is}^{-1}\Omega_{is} & M'_{is}\Pi_{is}^{-1}\Xi_{is}\\ \Omega'_{is}\Pi_{is}^{-1}M_{is} & \Omega'_{is}\Pi_{is}^{-1}\Omega_{is} & \Omega'_{is}\Pi_{is}^{-1}\Xi_{is}\\ \Xi'_{is}\Pi_{is}^{-1}M_{is} & \Xi'_{is}\Pi_{is}^{-1}\Omega_{is} & \Xi'_{is}\Pi_{is}^{-1}\Xi_{is} \end{pmatrix},
\end{aligned} \tag{11}$$

and we define the local $D$-criterion for the consider-then-choose model as the minimization of

$$\begin{aligned}
D^{\text{noncomp}}\text{-error} = \det\{I(\theta, \lambda_B, \zeta)^{-1}\}^{1/a}, \quad &\text{where}\\
a = q + m(m+1)/2 + R. &
\end{aligned} \tag{12}$$

It is important to note that the Fisher information matrix in (11) for a consider-then-choose model differs substantially from the information matrix in (6) for the compensatory model. This difference leads to very different designs, as will be shown through our comparative study in §5. Next, we describe the scenarios used for our comparative study and how the designs are constructed.

## 4. Design Scenarios and Construction

### 4.1. Design Scenarios

The design setting in which we conduct our comparative study involves four attributes, each with three levels. We follow Gilbride and Allenby (2006) to code the design matrix $X$ where the explanatory variables in $X$ for the no-choice option are set to zero. With the presence of the no-choice option, preferences for all three levels of one attribute (for example, brand) are estimable. For each of the remaining three attributes, only the two preference contrasts of the three levels are estimable. The preference parameter vector $\theta$ is therefore of length $3 + 3 \times 2 = 9$, the covariance matrix $\Lambda$ is of size $(9 \times 9)$, and the screening parameter vector $\zeta$ is of length $4 \times 3 = 12$. The covariates matrix $Z_i$ is set to the identity matrix for all respondents. Similar to Yu et al. (2009) and Sándor and Wedel (2002), the covariance matrix $\Lambda$ is set to the identity matrix $\mathrm{I}_9$ for the focal designs investigated in this paper. Other values of $\Lambda$, namely, $\Lambda = 2\mathrm{I}_9$, $\Lambda = 0.5\mathrm{I}_9$, and $\Lambda = 0.1\mathrm{I}_9$, are considered later when we investigate design robustness issues (see §6).

We consider two factors (preference and screening) each at two levels (high and low). This results in four

**Table 1** **Parameters for the Construction of Designs in the Comparative Study**

| Preference parameters $\theta_{9\times1}=\theta_0$ | Screening parameters $\zeta=\{\xi_1,\dots,\xi_{12}\}$ Low | High |
|---|---|---|
| Low | $\theta_0=0.25*(2,1,2,1,2,1,2,1,2)'$, $\zeta=\mathbf{1}_4\otimes(0.01,0.02,0.05)'$ | $\theta_0=0.25*(2,1,2,1,2,1,2,1,2)'$, $\zeta=\mathbf{1}_4\otimes(0.05,0.10,0.25)'$ |
| High | $\theta_0=(2,1,2,1,2,1,2,1,2)'$, $\zeta=\mathbf{1}_4\otimes(0.01,0.02,0.05)'$ | $\theta_0=(2,1,2,1,2,1,2,1,2)'$, $\zeta=\mathbf{1}_4\otimes(0.05,0.10,0.25)'$ |

*Note*. $\zeta=\mathbf{1}_4\otimes(0.01,0.02,0.05)'$ represents four copies of $(0.01,0.02,0.05)$ in vector $\zeta$; that is, $\zeta=(0.01,0.02,0.05,0.01,0.02,0.05,0.01,0.02,0.05,0.01,0.02,0.05)'$.

cells that are outlined in Table 1. We obtain optimal designs for compensatory and consider-then-choose models by minimizing the $D^{\text{comp}}$ error in (7) and the $D^{\text{noncomp}}$ error in (12), respectively. Note that only two scenarios (low and high preference) in Table 1 need to be considered for the designs constructed for the compensatory model because the screening parameters do not apply, whereas all four scenarios in Table 1 are considered for the designs constructed for the consider-then-choose model.

### 4.2. Design Construction Algorithm and Run Times

We use the coordinate-exchange algorithm proposed by Meyer and Nachtsheim (1995) to search for optimal designs. This algorithm has been found to be computationally efficient especially when a large number of attributes and attribute levels are involved (see Kessels et al. 2009 and Yu et al. 2009 for details). With coordinate-exchange, each attribute level in the initial design is exchanged with all possible levels of that attribute. A level change is accepted only if it improves the design criterion value. The algorithm starts the exchange at the first attribute of the first profile in the design. After it goes through the first round and finds the best exchanges for all the attributes of all profiles, it reiterates until no further substantial improvement is possible.

The evaluation of the Fisher information matrices (6) and (11) involves integrals. For a similar problem, Yu et al. (2009) demonstrated that the traditional Monte Carlo random sample is not computationally efficient for the evaluation of the integrals. They showed that more intelligent, quasi-random samples can achieve the same level of accuracy with a much smaller number of sampling points. Various quasi-random sampling methods have been proposed and applied in the marketing literature. These include the orthogonal array-based Latin hypercube sample used in Sándor and Wedel (2002) and the Halton sequences used in Yu et al. (2009). In this paper, we combine the modified Latin hypercube sample (MLHS) (see Hess et al. 2006) with recent advancements in orthogonal Latin hypercube designs (Sun et al. 2009) to achieve a more uniform coverage of the parameter space and enhanced performance in high dimensions. In our search of optimal designs, we found that a 128-point sample is sufficient. To mitigate the risk of the design search getting stuck in suboptima, we performed 100 runs for the search of optimal designs. A single run for the search of a local design (24 choice sets, 3 alternatives per choice set) for the setting in our comparative study took about 2.5 minutes for the compensatory model and about 10 minutes for the consider-then-choose model on a PC (2.66 GHz CPU and 3 GB RAM).

## 5. Comparative Study

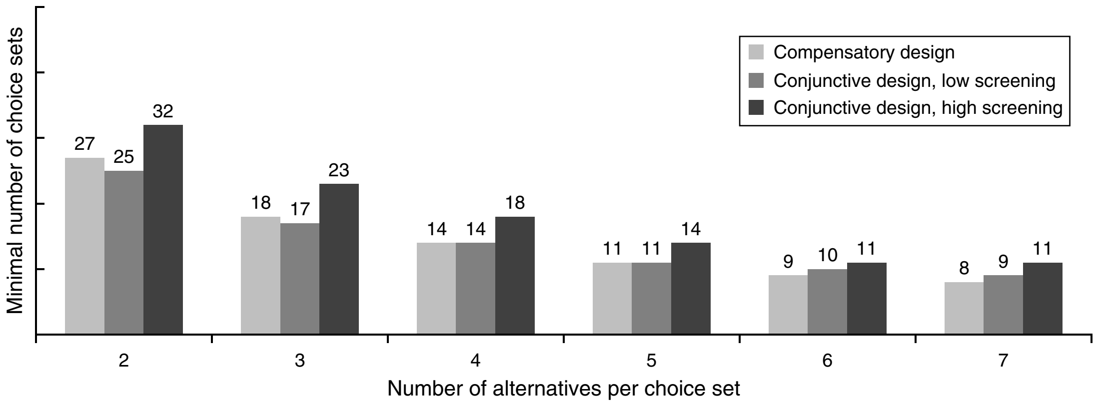
In this section, we contrast conjunctive designs with compensatory designs. As mentioned earlier, conjunctive designs are optimized for the conjunctive model. In particular, we examine the differences between these two design types on the following dimensions: size, characteristics, and relative performance. Note that we only consider full-profile designs where all attributes are present in each choice set and do not consider partial-profile designs where only a subset of attributes is present (see Green 1974).

### 5.1. Design Size: Are Conjunctive Designs Larger?

Are conjunctive designs larger than their compensatory counterparts? To answer this question, we determine the minimum number of choice sets required to estimate compensatory and conjunctive models given a fixed number of alternatives per choice set. The minimum number of choice sets for a compensatory model is determined based on the criterion that the information matrix be nonsingular. For the smallest compensatory design that satisfies this criterion, we measure the information it contains. Next, we look for the smallest conjunctive design that contains the same amount of information as the compensatory design. We explain our approach in greater detail next.

For a given number of alternatives per choice set, we start with a prespecified number of choice sets and randomly generate 100 designs in this setup. We evaluate these designs and examine whether there exists a design that leads to a nonsingular information matrix (Equation (6)). If the answer is yes, we reduce the number of choice sets. Conversely, if the answer is no, we increase the number of choice sets. With the updated number of choice sets in the setup, we generate 100 new designs and repeat our evaluation and examination process on these new designs. This continues until we find the minimum number of choice sets required for the compensatory design. Next, for this minimum number of choice sets with the given number of alternatives per choice set, we

**Figure 1** **Minimum Number of Choice Sets Required for the Low Preference Scenarios**




*Note.* The number of alternatives in the chart excludes the default "no-choice" alternative in each choice set.

identify the highest value of information—or the minimal value of the $D^{\text{comp}}$ error (Equation (7)) for the compensatory design—using the coordinate-exchange algorithm described in §4. We then use this value as our benchmark and apply the coordinate-exchange algorithm again to investigate the minimal number of choice sets required for the conjunctive designs to achieve the same amount of information. We conduct this exercise by varying the number of alternatives per choice set for each one of the four conditions listed in Table 1. We report in Figure 1 the results corresponding to the low preference conditions.

Note that there are a total of 54 hyperparameters (9 for $\theta + 9 \times 10/2$ for $\Lambda$) to be estimated for the compensatory model and 66 hyperparameters (9 for $\theta + 9 \times 10/2$ for $\Lambda + 12$ for $\zeta$) for the conjunctive model. A key finding in Figure 1 is that conjunctive designs do not require more choice sets than their compensatory counterparts require when the screening is low. However, when the screening is high, conjunctive designs do require more choice sets to achieve a comparable amount of information. Results for the high preference conditions are similar. For example, for the case of three alternatives per choice set, the minimum number of choice sets is 18 for the compensatory design, 17 for the conjunctive design with the low screening assumption, and 21 for the conjunctive design with the high screening assumption.

The results in Figure 1 make intuitive sense because information on the screening parameter for each attribute level is contained in the frequency with which it occurs in the chosen alternatives. These frequencies come "free" with the choice data, and therefore no additional choice sets are required for estimation of the screening parameters when the screening probabilities are low. However, when the screening probabilities are high, more respondents are likely to select the "no-choice" option. Such no-choice selections come at a cost—they do not provide any information on the preference parameters. To compensate for this information loss in situations where screening is high, conjunctive designs require more choice sets than compensatory designs.

Further examination of the results in Figure 1 reveals that the use of three to five alternatives per choice set is appropriate for typical marketing applications. A number less than three will remarkably increase the number of choice sets required, and a number more than five may require considerable cognitive resources for the simultaneous comparison of multiple alternatives. Guided by results in this section, for the remainder of this paper, we use a design specification with 24 choice sets and three alternatives per choice set for all our comparisons between compensatory and conjunctive designs.

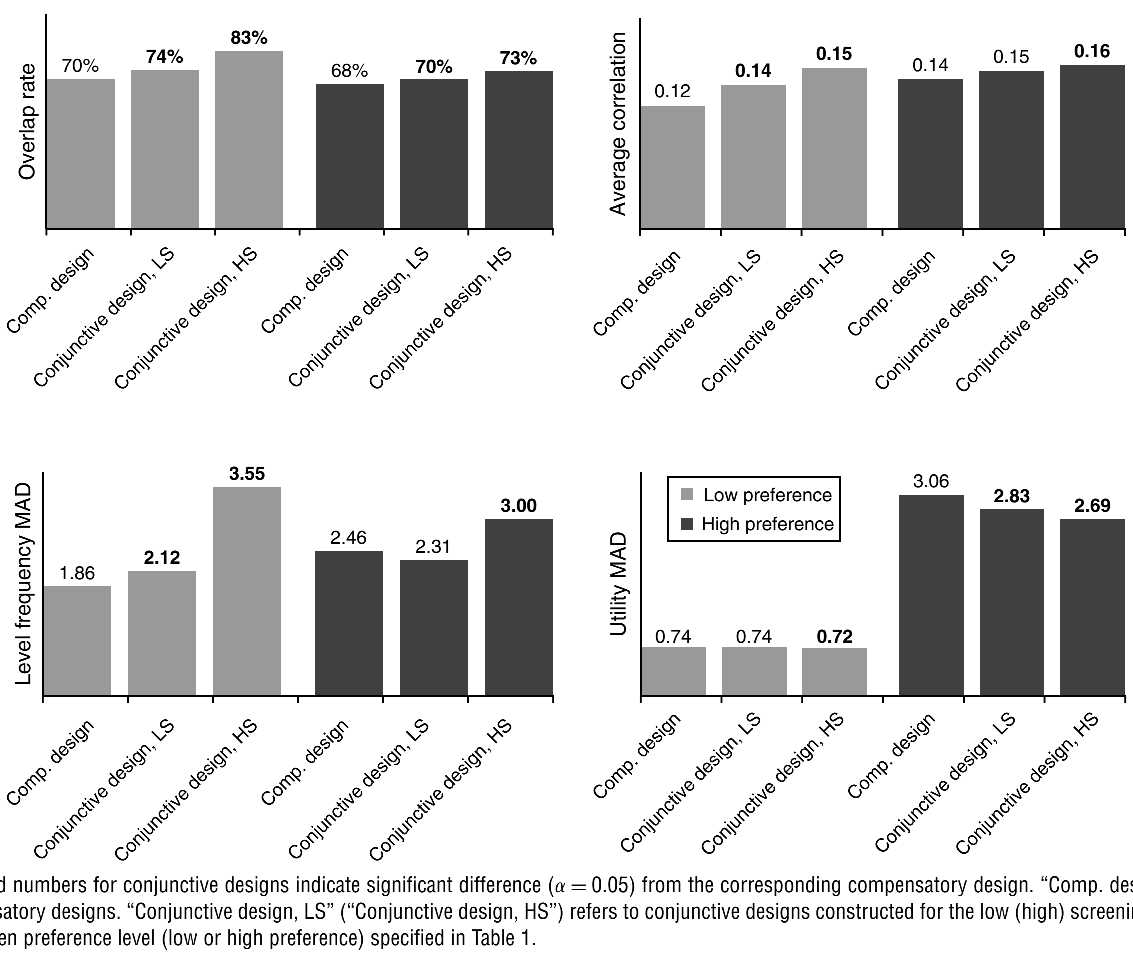
### 5.2. Design Characteristics: Do Conjunctive Designs Have Different Properties?

Extant research on compensatory choice designs has emphasized the following four design aspects: level overlap, orthogonality, level balance, and utility balance. Next, we study how conjunctive designs differ from compensatory designs on these four aspects. For each scenario in Table 1, we apply the coordinate-exchange algorithm to construct efficient compensatory and conjunctive designs. For 100 randomly generated initial designs, we obtain 100 sets of efficient compensatory and conjunctive designs and compare the characteristics of these designs on four dimensions: level overlap, orthogonality, level balance, and utility balance.

Level overlap of an attribute refers to the situation when there is a repeated occurrence of an attribute level within a choice set. We follow the method used in Sándor and Wedel (2002) to measure the percentage of level overlap in a design. For our setting of three alternatives per choice set, we calculate the percentage of columns in the choice sets that have repeated occurrences of attribute levels. The higher this percentage, the higher the level overlap. We note that although this standard measure of level overlap works well for the purpose of our study, it ignores overlap intensity. For example, it does not distinguish an attribute level overlap that occurs in two alternatives of a choice set from an attribute level overlap that occurs in three alternatives. To measure orthogonality, we use the average pairwise correlation of the mean-adjusted column vectors in model matrix $X$. An orthogonal design has zero correlation. The farther away from zero the average correlation, the less orthogonal is the design. To measure level balance, we use the mean absolute deviation (MAD) of level frequency across attribute levels. Absolute deviation of each attribute level is the difference of its frequency in the design from its expected frequency, which equals 24 for each attribute level in our design setting. A design with a MAD of zero is level balanced, with equal occurrence of each attribute level within an attribute. Larger MAD implies lower level balance. Similarly, we use utility MAD to measure utility balance. The utilities are calculated as $X\theta_0$ (see values of $\theta_0$ in Table 1). A design with a utility MAD of zero is utility balanced, with equally attractive alternatives. Larger MAD implies lower utility balance.

Figure 2 presents how the four characteristics between the two design types differ. For each characteristic, mean statistics across the 100 designs are reported. For each preference scenario, we use $t$-tests to assess whether there is significant difference between the conjunctive design and the corresponding compensatory design. Bold numbers in Figure 2 for the conjunctive designs indicate significant difference ($\alpha = 0.05$) from the corresponding benchmark compensatory designs. For example, for the low preference scenario, the overlap rate is 74% for the conjunctive design constructed under the low screening assumption and 83% for the conjunctive design constructed under the high screening assumption. Both are significantly different from the 70% overlap rate for the benchmark compensatory design. The results in Figure 2 suggest that on average the conjunctive designs have higher level overlap, lower orthogonality, less level balance, and more utility balance than the corresponding compensatory designs. Differences between compensatory and conjunctive designs

**Figure 2** **Design Characteristics Comparison**




*Notes.* Bold numbers for conjunctive designs indicate significant difference ($\alpha = 0.05$) from the corresponding compensatory design. “Comp. design” refers to compensatory designs. “Conjunctive design, LS” (“Conjunctive design, HS”) refers to conjunctive designs constructed for the low (high) screening scenario and the given preference level (low or high preference) specified in Table 1.

in all four measures are significantly amplified at higher screening levels: overlap rate and utility balance increase, whereas orthogonality and level balance decrease as screening levels increase. Because of the repetitions of the same attribute level within a choice set when there is level overlap, it makes intuitive sense that higher overlap rate is associated with higher utility balance but lower orthogonality and lower level balance. Examples of compensatory and conjunctive designs with their respective level overlap rates are provided in Appendix C.

There is a good reason why conjunctive designs have higher level overlap than corresponding compensatory designs. Level overlap helps detect and pinpoint screening, and thus designs with higher overlap likely have higher efficiency in estimating the screening parameters. To help understand this, consider a stylized example with two attributes $A$ and $B$, each with two levels. For ease of illustration, let us assume that, on average, the attributes are equally important and levels within each attribute are equally attractive—that is, all elements of $\theta_0$ are the same. Consider a choice set with two alternatives with no level overlap, ($A1B2$, $A2B1$). If one alternative, say, $A1B2$, is chosen significantly less frequently than 50% of the time, then there is indication of some screening in the sample. However, without level overlap in the choice set, we cannot tell whether the screening is on $A1$, the first level of the attribute $A$, or $B2$, the second level of attribute $B$. On the other hand, consider a choice set with level overlap on $A1$, that is, a choice set with alternatives ($A1B2$, $A1B1$). For this choice set ($A1B2$, $A1B1$), screening on $A1$ will manifest itself via the predominant selection of the no-choice option. In contrast, screening on $B2$ will show up with fewer respondents choosing $A1B2$. Clearly, data from the choice set with level overlap on $A1$ (i.e., $A1B2$, $A1B1$) are more informative about where screening occurs.

### 5.3. Using Compensatory Designs When the True Model Is Conjunctive: What Are the Losses?

In the absence of any direction on how to design an experiment for a conjunctive model, compensatory designs offer the only practical alternative. In this section, we investigate the design performance of compensatory designs when the true model is conjunctive. We use three measures: design efficiency loss based on the $D^{\text{noncomp}}$-error in (12), design efficiency loss based on parameter recovery, and design power in model detection.

**Efficiency Loss Based on the $D^{\text{noncomp}}$-Error.** For a given level of preference (e.g., low preference) we take the best compensatory design from our search and evaluate at a given screening level (e.g., low screening) the efficiency loss of this design based on the $D^{\text{noncomp}}$-error. The efficiency loss of a design $X$ under the scenario $\Omega_{ij}$ ($i =$ low preference, high preference; $j =$ low screening, high screening) is defined as

$$\text{Efficiency loss based on } D^{\text{noncomp}}\text{-error}(X \mid \Omega_{ij}) = 1 - \frac{D^{\text{noncomp}}\text{-error}(X^* \mid \Omega_{ij})}{D^{\text{noncomp}}\text{-error}(X \mid \Omega_{ij})}, \tag{13}$$

where $X^*$ is the corresponding optimal conjunctive design from the computer search.

**Efficiency Loss Based on Errors in Parameter Recovery.** In addition to the efficiency measure (13) based on the $D^{\text{noncomp}}$-error, we also examine through simulation the efficiency loss based on the expected root mean square errors (ERMSEs) of the posterior estimates of the preference and screening parameters $\theta$ and $\zeta$. In particular, the ERMSEs of the posterior estimates of $\theta$ and $\zeta$ under a given design $X$ are defined respectively as

$$\text{ERMSE}(\theta \mid X) = \int [(\hat{\theta} - \theta)'(\hat{\theta} - \theta)]^{1/2} f(\hat{\theta} \mid X)\, d\hat{\theta}, \quad \text{and}$$

$$\text{ERMSE}(\zeta \mid X) = \int [(\hat{\zeta} - \zeta)'(\hat{\zeta} - \zeta)]^{1/2} f(\hat{\zeta} \mid X)\, d\hat{\zeta},$$

where $f(\hat{\theta} \mid X)$ and $f(\hat{\zeta} \mid X)$ denote the distributions of the posterior estimates $\hat{\theta}$ and $\hat{\zeta}$ from data generated by using the design $X$ and the true parameter values $\theta$ and $\zeta$. The posterior estimates $\hat{\theta}$ and $\hat{\zeta}$ are obtained by using Markov chain Monte Carlo with random-walk Metropolis-Hastings steps (see Rossi et al. 2005, Chapter 3). For a given design we simulate responses from 1,000 respondents based on the conjunctive model and obtain the ERMSEs for each scenario in Table 1. The higher the ERMSE, the less efficient the corresponding design is in parameter recovery. The efficiency loss based on parameter recovery is defined the same way as in (13) by replacing the $D^{\text{noncomp}}$-errors with the ERMSEs of the corresponding parameter ($\theta$ or $\zeta$).

**Power in Model Detection.** For each scenario in our simulation, we fit two models to the simulated data—the conjunctive model and the compensatory model. We first assess the model fit using log marginal density (LMD) calculated according to the mean harmonic estimator by Newton and Raftery (1994). Noting that the LMD can be biased in favor of more complex models, we also calculate the deviance information criterion (DIC), which involves a penalty term for model complexity. The DIC measure for model fit was proposed by Spiegelhalter et al. (2002) specifically for hierarchical models. Note that the higher the LMD, the better the model fit, whereas the lower the DIC, the better the model fit. Therefore, for the data generated by using design $X$, we compute two difference measures: "LMD (for the conjunctive

**Table 2** Efficiency Loss from Using a Compensatory Design When the True Model Is Conjunctive

| Screening level | Preference level | Percentage efficiency loss based on $D^{\text{noncomp}}$ | ERMSE($\theta$) | ERMSE($\zeta$) |
|---|---|---|---|---|
| Low | Low | 25.9 | 2.1 | 18.3 |
| | High | 28.5 | 19.2 | 47.2 |
| High | Low | 33.0 | 21.2 | 45.0 |
| | High | 27.0 | 16.0 | 29.2 |

model) − LMD (for the compensatory model)" and "DIC (for the compensatory model) − DIC (for the conjunctive model)." The higher the difference measures, the more power the design $X$ has in detecting that the true model of the choice decision process is conjunctive.

For each scenario in Table 1, we report the measures of efficiency loss in Table 2. Based on the $D^{\text{noncomp}}$-error, we find that the compensatory designs have significant efficiency loss ranging from 25.9% to 33% over the corresponding conjunctive designs. Efficiency losses based on the ERMSEs of the posterior estimates of $\theta$ or $\zeta$ are also large, confirming that compensatory designs do not work well for the conjunctive model. These findings make a strong case for the use of conjunctive designs in order to achieve higher design efficiency.

For each scenario in Table 1, we report the measures of power in model detection in Table 3. Recall that the higher the difference measures, the more power the design has in detecting the true conjunctive model. Across all four conditions, the DIC difference measures for conjunctive designs are larger than those for compensatory designs. An identical pattern of results holds true for the LMD difference measure. Based on both DIC and LMD difference measures, our results suggest that conjunctive designs have higher power than compensatory designs in detecting the true underlying conjunctive model. Although expected, it is instructive to contrast this result with what we investigate next.

### 5.4. Using Conjunctive Designs When the True Model Is Compensatory: What Are the Losses?

The results in §5.3 provide strong evidence in support of conjunctive designs when the underlying decision process is conjunctive. In this section we investigate the converse: how well do conjunctive designs perform when the true choice decision process is compensatory? Naturally, there will be some loss of efficiency from using a conjunctive design. An interesting question is whether this efficiency loss will be as big as what we saw in the previous section. In addition, how does a conjunctive design compare to a compensatory design in its ability to detect a true compensatory decision process?

We use the same measures as presented in §5.3, except that the efficiency loss is now based on the $D^{\text{comp}}$-error rather than the $D^{\text{noncomp}}$-error. Specifically, the efficiency loss of a design $X$ under the scenario $\Omega_i$ ($i =$ low preference, high preference) is defined as

$$\text{Efficiency loss based on } D^{\text{comp}}\text{-error}(X \mid \Omega_i) = 1 - \frac{D^{\text{comp}}\text{-error}(X^* \mid \Omega_i)}{D^{\text{comp}}\text{-error}(X \mid \Omega_i)}, \tag{14}$$

where $X^*$ is the corresponding optimal compensatory design.

Results reported in Table 4 show that the efficiency losses of conjunctive designs are relatively low when the true model is compensatory. In particular, the range of efficiency losses is 2.3% to 13% based on the $D^{\text{comp}}$-error. This is much lower than the range of efficiency losses of compensatory designs (25.9% to 33%) under the conjunctive model as reported in Table 2. The efficiency losses of conjunctive designs are especially low (2.3% to 5.5%) when low screening is assumed in the design construction.

Results pertaining to power in model detection, reported in Table 5, are also quite different from the previous section. All difference measures for DIC and LMD are very close to zero. That is, both conjunctive designs and compensatory designs have similar power in model detection when the true underlying model is compensatory. This is in contrast with results

**Table 3** Measures of Power in Model Detection When the True Model Is Conjunctive

| | | LMD difference (Conjunctive − Compensatory) | | DIC difference (Compensatory − Conjunctive) | |
|---|---|---|---|---|---|
| Screening level | Preference level | Compensatory design | Conjunctive design | Compensatory design | Conjunctive design |
| Low | Low | 4.6 | 10.7 | 12.0 | 15.2 |
| | High | 64.5 | 138.9 | 121.3 | 266.1 |
| High | Low | 37.5 | 52.8 | 75.9 | 97.8 |
| | High | 181.1 | 289.6 | 358.5 | 570.4 |

**Table 4 Efficiency Loss from Using Conjunctive Designs When the True Model Is Compensatory**

| | Conjunctive designs with low screening assumed | | Conjunctive designs with high screening assumed | |
|---|---|---|---|---|
| | Percentage efficiency loss based on | | | |
| Preference level | $D^{comp}$ | ERMSE($\theta$) | $D^{comp}$ | ERMSE($\theta$) |
| Low | 2.3 | 3.8 | 13.0 | 10.0 |
| High | 5.5 | 4.4 | 7.3 | 1.7 |

in Table 3 that suggest that compensatory designs have lower power than conjunctive designs when the true model is conjunctive. The main lesson is that regardless of whether the true underlying model is compensatory or conjunctive, it is safer to use the conjunctive design.

## 6. Robustness of Conjunctive Designs

Our findings from the comparative study in §5 make a strong case for the use of conjunctive designs. In this section, we investigate the robustness of conjunctive designs when the assumed values of the covariance matrix, preference, and screening levels deviate from the true values. We report the mean efficiency loss measure based on the $D^{\text{noncomp}}$-error together with the minimum and the maximum efficiency losses (in brackets) obtained through resampling the standard normal errors 100 times in the evaluation of the integrals in the Fisher information matrix (11).

Table 6 shows that conjunctive designs are fairly robust to misspecifications of the random effects covariance matrix $\Lambda$. For each preference and screening scenario, we take the conjunctive design constructed with $\Lambda = \mathrm{I}$ and check its efficiency loss over the corresponding optimal conjunctive design constructed with the true covariance matrix, $\Lambda = 2\mathrm{I}$, 0.5I, or 0.1I. In general, the efficiency losses are small. Only when the assumed covariance matrix ($\Lambda = \mathrm{I}$) is 10 times the true covariance matrix ($\Lambda = 0.1\mathrm{I}$) does the efficiency loss rise to the neighborhood of 10%.

Robustness checks for misspecification of preference and screening parameters (see Table 7) reveal that the efficiency losses are higher. Specifically, the efficiency loss can be as high as 20% when both the preference and screening parameters used in the design construction deviate from the true parameters. For example, when the true screening is low and preference is high, the conjunctive design constructed by assuming high screening and low preference has an average efficiency loss of 19.8%. On the one hand, when the assumed preference is consistent with the true preference but the assumed screening deviates from the true screening, the efficiency loss of the corresponding conjunctive design is low (3.0% to 8.3%). On the other hand, when the assumed screening is consistent with the true screening but the assumed preference deviates from the true preference, the efficiency loss is higher (10.1% to 18.1%). This suggests that the conjunctive designs are more sensitive to misspecification of the preference parameters than they are to that of the screening parameters. Therefore, it is important to incorporate prior knowledge, especially

**Table 5 Measures of Power in Model Detection When the True Model Is Compensatory**

| | LMD difference (Conjunctive − Compensatory) | | | DIC difference (Compensatory − Conjunctive) | | |
|---|---|---|---|---|---|---|
| | Compensatory design | Conjunctive design | | Compensatory design | Conjunctive design | |
| | | Screening level | | | Screening level | |
| Preference level | | Low | High | | Low | High |
| Low | 0.18 | −1.2 | −0.2 | −0.8 | 0.7 | −2.80 |
| High | −0.03 | 1.5 | 1.7 | 0.2 | −3.8 | 0.08 |

**Table 6 $D^{\text{noncomp}}$-Error Efficiency Loss Because of Covariance Misspecification**

| Assumed model parameters in design construction | | True covariance matrix | | |
|---|---|---|---|---|
| Screening level | Preference level | $\Lambda = 2\mathrm{I}$ | $\Lambda = 0.5\mathrm{I}$ | $\Lambda = 0.1\mathrm{I}$ |
| Covariance matrix $\Lambda = \mathrm{I}$ | | | | |
| Low | Low | 0.8 [−0.8, 5.4] | 2.4 [0.1, 4.8] | 5.6 [4.3, 7.0] |
| | High | 6.0 [1.0, 9.3] | 2.1 [−2.6, 4.7] | 11.2 [9.3, 13.1] |
| High | Low | 0.6 [−3.4, 3.8] | 1.7 [−0.1, 4.4] | 9.0 [7.5, 10.1] |
| | High | 1.6 [−1.2, 6.7] | 0 [0.0, 0.0] | 11.3 [8.2, 13.0] |

**Table 7 $D^{\text{noncomp}}$-Error Efficiency Loss Because of Preference and Screening Misspecification**

| Assumed model parameters in design construction | | True model parameters $\Lambda = \mathrm{I}$ | | | |
|---|---|---|---|---|---|
| | | Low screening | | | High screening |
| Screening level | Preference level | Low preference | High preference | Low preference | High preference |
| Random effects covariance matrix $\Lambda = \mathrm{I}$ | | | | | |
| Low | Low | 0 | 13.1 [10.1, 15.0] | 7.8 [4.6, 10.3] | 14.5 [10.3, 16.7] |
| | High | 10.6 [6.2, 12.6] | 0 | 19.7 [16.1, 23.1] | 4.6 [1.2, 7.4] |
| High | Low | 8.3 [6.2, 11.4] | 19.8 [17.1, 23.3] | 0 | 10.1 [1.7, 12.9] |
| | High | 12.6 [10.9, 14.0] | 3.0 [0.4, 6.2] | 18.1 [15.8, 19.9] | 0 |

about preference parameters, when constructing conjunctive designs.

## 7. Summary and Conclusions

Existing literature in marketing offers good direction on how to construct efficient designs for compensatory choice models. In contrast, little research exists on how to construct efficient designs for consider-then-choose models. Our research fills this void and provides a method to construct efficient designs for a consider-then-choose model encompassing disjunctive, conjunctive, subset conjunctive, and disjunctions of conjunctions screening rules.

We have focused on local designs in this paper. Our findings make a strong case for the use of conjunctive designs when there is prior evidence to support respondent screening. We find that conjunctive designs are quite different from compensatory designs—they are characterized by a higher level overlap. Higher overlap also results in lower orthogonality, less level balance, and more utility balance in conjunctive designs. We explain why higher level overlap may be necessary to pinpoint screening behavior. This is an important finding because it suggests that readily available orthogonal or nearly orthogonal designs that have minimal level overlap (Huber and Zwerina 1996) should not be used with consider-then-choose models that involve noncompensatory screening. Practitioners appear to be aware of this design issue because they note that that minimal overlap designs do not work well when respondents have "must-have" or "must-avoid" features in the choice process (e.g., Johnson 2008, Orme 2009). Unlike ad hoc approaches such as the "balanced overlap" (Orme 2009) adopted by practitioners, we formalize the process of creating efficient designs for models that incorporate consideration sets. The statistical machinery we provide helps create designs that permit efficient estimation of both preference and screening parameters in a consider-then-choose model.

In terms of design performance, we find that compensatory designs are much less efficient than corresponding conjunctive designs when the true underlying choice model is conjunctive. In contrast, when the true underlying choice process is compensatory, the efficiency loss of using conjunctive designs over compensatory designs is rather small. It is noteworthy that we find that when the true choice process is unknown, conjunctive designs may be the safer bet—they have about the same power as compensatory designs in identifying a true compensatory choice process and much higher power when the choice process is conjunctive. This finding strengthens the case for the use of conjunctive designs because in practice, the true choice process is unknown.

It is reasonable to believe that a consideration set resulting from disjunctive screening is likely more similar to that of the compensatory model than to that of the conjunctive model. A disjunctive screening ($k = 1$) could be viewed as a fairly mild rule that excludes few alternatives from consideration, thus appearing more like a compensatory model. In contrast, a consideration set resulting from the subset conjunctive screening rule with a relatively high subset size $k$ is likely to be more similar to that of the conjunctive model than to that of the compensatory model. In other words, disjunctive and subset conjunctive screening models could be viewed as lying between the two extremes characterized by the compensatory and conjunctive models. Therefore, we expect that a conjunctive design will likely perform worse than a compensatory design when the true screening process is disjunctive. Similarly, a compensatory design will likely perform worse than a conjunctive design when the true data-generating mechanism involves subset conjunctive screening. On the other hand, a conjunctive design likely performs better than a compensatory design when the true screening process is subset conjunctive with a relatively high subset size $k$, and a compensatory design likely performs better than a conjunctive design when the true data-generating mechanism involves disjunctive screening. Design efficiency calculations that verify this somewhat intuitive result are available from the authors upon request. The key insight from the above discussion is this: it is critical that the decision to use a conjunctive, disjunctive, compensatory, or any other design type be based on reliable prior information.

## 8. Future Research

This paper sets the stage for important follow-up research in the area of choice designs for consider-then-choose models in particular and noncompensatory models in general. This includes research ideas that are related to our paper but beyond the scope of our current investigation. We end the paper with a brief discussion of these topics.

### 8.1. Prior Elicitation

A practical issue to consider in design construction is specification of the preference and screening parameters of the consider-then-choose model. We find it important to incorporate prior knowledge, especially the magnitude of preference parameters, in the construction of consider-then-choose designs to improve design efficiency. Such knowledge can be obtained through pretests, self-explicated questions, or managers' beliefs based on past data. A careful investigation of the best approach to obtain prior information on preference and screening parameters is necessary for widespread adoption of these designs. An attractive aspect of conjunctive designs is that self-explicated data on a must-have or must-avoid feature is easy to obtain and may serve as a reasonable prior for the screening parameters in conjunctive designs. Because the conjunctive designs are sensitive to prior misspecification, it may also be helpful to obtain measures of uncertainty around such self-explicated data (Sándor and Wedel 2001).

### 8.2. Independence of Screening and Choice

In the consider-then-choose model we currently use, screening and preference parameters are assumed to be independent. A possible extension is to model the correlation between the two stages of screening and choice through common covariates. For example, both the preference and screening parameters associated with a respondent can be modeled as functions of covariates that capture information of the respondent (for example, gender and age group). Our current design machinery can be adapted to such a model without too much difficulty. An alternative approach to modeling correlation between the two stages is by employing a "compensatory screening rule" that excludes an alternative from consideration if the utility of the alternative is below a threshold. In such models the correlation between the screening and choice stages is captured through common variables in the two utilities—the screening utility can be the same as the choice utility (e.g., Jedidi et al. 1996) or different (e.g., van Nierop et al. 2010). For such models, the search for optimal designs becomes more complex because the choice probability cannot always be expressed in a closed form. This makes the Fisher information matrix intractable. Future research is required to provide guidance on designing choice experiments for consider-then-choose models that use a compensatory screening rule. Along the same lines it would be worthwhile to study design criteria for lexicographic (e.g., Kohli and Jedidi 2007) or elimination by aspects (e.g., Gilbride and Allenby 2006) screening rules that may or may not follow a two-stage choice process.

### 8.3. Design (and Estimation) Algorithms for Subset Conjunctive and Disjunctions of Conjunctions Models

Consider-then-choose models that use subset conjunctive or disjunctions of conjunctions rules present additional design challenges. For a subset conjunctive rule, prior knowledge of subset size $k$ is hard to assess. The problem is particularly complicated if subset size $k$ is heterogeneous across individuals. A similar problem exists for the disjunctions of conjunctions rule. The patterns for the disjunctions of conjunctions rule are also difficult to ascertain because of a large number of possible combinations and heterogeneity across individuals. Although our proposed design approach applies for a given subset size $k$ (subset conjunctive) or a given set of patterns (disjunctions of conjunctions), this is somewhat restrictive. Additional work is required to extend our approach to heterogeneous subset sizes or patterns that follow a probability distribution. The logical first step in that direction would be to develop hierarchical Bayes algorithms to estimate the consider-then-choose models where the first stage involves subset conjunctive (disjunctions of conjunctions) screening rules with heterogeneous subset sizes (patterns). To the best of our knowledge, such models do not currently exist. Jedidi and Kohli (2005) and Hauser et al. (2010) provide excellent starting points for such an endeavor.

### 8.4. Heterogeneous Designs

Our focus in this paper was on homogeneous designs where each respondent gets the same design. Heterogeneous designs where different respondents, or groups of respondents, are given different designs can be constructed to achieve higher efficiency (Sándor and Wedel 2005). Alternatively, a big design can be split into multiple versions of small designs and be administered as heterogeneous designs to different respondents. Heterogeneous designs may be necessary for contexts where a large number of attributes and levels are involved. In addition, they could be used when respondents belong to distinct preference/screening groups and could therefore benefit from customized designs. Existing approaches to heterogeneous designs involve random allocation of different designs to different individuals. An interesting topic for future research is to find a systematic approach that effectively incorporates information on covariates ($Z_i$) to customize designs for individuals belonging to different groups.

### 8.5. Alternative Design Criteria

Although we have focused on the $D$-error criterion in this paper, the framework can be easily extended to other design criteria that are based on the same information matrix (11). For example, it can be extended to the $A$-criterion by minimizing the sum of the diagonal elements of the inverse of the Fisher information matrix (11). Similarly, it can be extended to the $V$- or $G$-criteria when the interest is in future predictions (see Kessels et al. 2006) or a managerially relevant criterion when the interest is in the estimation of linear or nonlinear functions of model parameters such as the estimation of willingness to pay (see Toubia and Hauser 2007). The impact of alternative design criteria on design properties is an important area that deserves future research.

### 8.6. Semi-Bayesian or Full-Bayesian Designs

The local design criterion we develop in this paper can be easily extended to the semi-Bayesian or the full-Bayesian design criterion. By formally incorporating uncertainty of the hyperparameters, these Bayesian designs are likely to be more efficient than local designs. However, a major hurdle for such an extension is that it is computationally impractical. For the setting used in our comparative study, we found that it would take a PC (2.66 GHz CPU and 3 GB RAM) more than three months to complete the 100 runs of search for an optimal semi-Bayesian design. It would take even longer for the search of an optimal full-Bayesian design. The computation time increases exponentially as the number of attributes and levels increases. Alternative algorithms that improve the computational speed for the search of semi-Bayesian or full-Bayesian designs would be a useful addition to the marketing researcher's tool kit.

## Appendix A. Information Matrix for the Compensatory Model

In this appendix, we show the derivation of the information matrix $I(\theta, \lambda_B)$ for the compensatory model where the random effects may be correlated, and the $m(m+1)/2$ unique parameters in the general-form covariance matrix $\Lambda$ is denoted by vector $\lambda_B$.

The likelihood function of $\theta$ and $\Lambda$ is $L(\theta,\Lambda) = f(y \mid \theta, \Lambda) = \prod_{i=1}^N \prod_{s=1}^S f(y_{is} \mid \theta, \Lambda)$.

The Fisher information matrix $I(\theta, \lambda_B)$ can be expressed as

$$I(\theta, \lambda_B) = \begin{bmatrix} -E\dfrac{\partial^2}{\partial\theta\,\partial\theta'}\log L(\theta,\Lambda) & -E\dfrac{\partial^2}{\partial\theta\,\partial\lambda_B'}\log L(\theta,\Lambda) \\ -E\dfrac{\partial^2}{\partial\lambda_B\partial\theta'}\log L(\theta,\Lambda) & -E\dfrac{\partial^2}{\partial\lambda_B\partial\lambda_B'}\log L(\theta,\Lambda) \end{bmatrix} \quad \text{(A1)}$$

Note that

$$E\left[\frac{\partial^2}{\partial\theta\,\partial\theta'}\log f(y \mid \theta,\Lambda)\right] = E\left[\frac{\partial^2}{\partial\theta\,\partial\theta'}\sum_{i=1}^N\sum_{s=1}^S \log f(y_{is} \mid \theta,\Lambda)\right] = \sum_{i=1}^N\sum_{s=1}^S E\left[\frac{\partial^2}{\partial\theta\,\partial\theta'}\log f(y_{is} \mid \theta,\Lambda)\right],$$

and

$$\frac{\partial^2}{\partial\theta\partial\theta'}\log f(y_{is}|\theta,\Lambda) = \frac{\partial}{\partial\theta}\left\{\frac{\partial f(y_{is}|\theta,\Lambda)/\partial\theta'}{f(y_{is}|\theta,\Lambda)}\right\} = \frac{\partial^2 f(y_{is}|\theta,\Lambda)/\partial\theta\partial\theta'}{f(y_{is}|\theta,\Lambda)} - \frac{(\partial f(y_{is}|\theta,\Lambda)/\partial\theta)\cdot(\partial f(y_{is}|\theta,\Lambda)/\partial\theta')}{f(y_{is}|\theta,\Lambda)f(y_{is}|\theta,\Lambda)}.$$

Since

$$\mathrm{E}\left[\frac{(\partial^2 f(y_{is} \mid \theta,\Lambda)/\partial\theta\,\partial\theta')}{f(y_{is} \mid \theta,\Lambda)}\right] = \frac{\partial^2}{\partial\theta\,\partial\theta'}\int f(y_{is} \mid \theta,\Lambda)\,dy_{is} = \frac{\partial^2}{\partial\theta\,\partial\theta'}(1) = 0,$$

$$E\frac{\partial^2}{\partial\theta\,\partial\theta'}\log L(\theta,\Lambda) = -\sum_{i=1}^N\sum_{s=1}^S E\left[\frac{(\partial f(y_{is}|\theta,\Lambda)/\partial\theta)\cdot(\partial f(y_{is}|\theta,\Lambda)/\partial\theta')}{f(y_{is}|\theta,\Lambda)f(y_{is}|\theta,\Lambda)}\right] = -\sum_{i=1}^N\sum_{s=1}^S\sum_{j=1}^J \frac{(\partial f(y_{is}=j|\theta,\Lambda)/\partial\theta)\cdot(\partial f(y_{is}=j|\theta,\Lambda)/\partial\theta')}{f(y_{is}=j|\theta,\Lambda)}. \quad \text{(A2)}$$

Similarly,

$$E\frac{\partial^2}{\partial\theta\partial\lambda_B'}\log L(\theta,\Lambda) = -\sum_{i=1}^N\sum_{s=1}^S\sum_{j=1}^J \frac{(\partial f(y_{is}=j \mid \theta,\Lambda)/\partial\theta)\cdot(\partial f(y_{is}=j \mid \theta,\Lambda)/\partial\lambda_B')}{f(y_{is}=j \mid \theta,\Lambda)}, \quad \text{(A3)}$$

$$E\frac{\partial^2}{\partial\lambda_B\partial\lambda_B'}\log L(\theta,\Lambda) = -\sum_{i=1}^N\sum_{s=1}^S\sum_{j=1}^J \frac{(\partial f(y_{is}=j \mid \theta,\Lambda)/\partial\lambda_B)\cdot(\partial f(y_{is}=j \mid \theta,\Lambda)/\partial\lambda_B')}{f(y_{is}=j \mid \theta,\Lambda)}, \quad \text{(A4)}$$

where $f(y_{is}=j \mid \theta,\Lambda)$ is the probability mass function, and

$$f(y_{is}=j \mid \theta,\Lambda) = \int f(y_{is}=j \mid \theta,\varepsilon) f(\varepsilon \mid \Lambda)d\varepsilon = \int \frac{\exp\{x_{sj}'(Z_i\theta+\varepsilon)\}}{\sum_{j=1}^J \exp\{x_{sj}'(Z_i\theta+\varepsilon)\}} f(\varepsilon \mid \Lambda)\,d\varepsilon, \quad \text{(A5)}$$

with $\varepsilon \sim$ Multivariate Normal(0, $\Lambda$). Let

$$p_{isj}(\varepsilon) = \frac{\exp\{x_{sj}'(Z_i\theta+\varepsilon)\}}{\sum_{j=1}^J \exp\{x_{sj}'(Z_i\theta+\varepsilon)\}};$$

then

$$\frac{\partial}{\partial\theta} f(y_{is}=j \mid \theta,\Lambda) = \int \frac{\partial}{\partial\theta}p_{isj}(\varepsilon)f(\varepsilon \mid \Lambda)\,d\varepsilon = \int \left\{p_{isj}(\varepsilon)Z_i'x_{sj} - p_{isj}(\varepsilon)\sum_{j=1}^J p_{isj}(\varepsilon)Z_i'x_{sj}\right\} f(\varepsilon \mid \Lambda)\,d\varepsilon. \quad \text{(A6)}$$

$$\frac{\partial}{\partial\lambda_B} f(y_{is}=j \mid \theta,\Lambda) = \int \frac{\partial}{\partial\lambda_B}p_{isj}(\varepsilon)f(\varepsilon \mid \Lambda)\,d\varepsilon = \int p_{isj}(\varepsilon)\frac{\partial}{\partial\lambda_B}f(\varepsilon \mid \Lambda)\,d\varepsilon. \quad \text{(A7)}$$

Decompose $\lambda_B$ into the $m$ diagonal elements of $\Lambda$, $\lambda_{rr}(r=1,\ldots,m)$, and the $m(m-1)/2$ upper off-diagonal elements of $\Lambda$, $\lambda_{rt}(r<t,\, r,t=1,\ldots,m)$; we then have

$$\frac{\partial}{\partial\lambda_{rr}} f(\varepsilon \mid \Lambda) = \frac{1}{(2\pi)^{m/2}}\left\{\frac{-[\Lambda^{-1}]_{rr}}{2\mid\Lambda\mid^{1/2}}\exp\left(-\frac{1}{2}\varepsilon'\Lambda^{-1}\varepsilon\right) + \frac{1}{2\mid\Lambda\mid^{1/2}}[\Lambda^{-1}\varepsilon\varepsilon'\Lambda^{-1}]_{rr}\exp\left(-\frac{1}{2}\varepsilon'\Lambda^{-1}\varepsilon\right)\right\}$$

$$= f(\varepsilon \mid \Lambda)\left\{\frac{1}{2}[\Lambda^{-1}\varepsilon\varepsilon'\Lambda^{-1}]_{rr} - \frac{1}{2}[\Lambda^{-1}]_{rr}\right\}, \quad \text{(A8)}$$

$$\frac{\partial}{\partial \lambda_{rt}} f(\varepsilon \mid \Lambda) = \frac{1}{(2\pi)^{m/2}}\left\{\frac{-[\Lambda^{-1}]_{rt}}{|\Lambda|^{1/2}}\exp\left(-\frac{1}{2}\varepsilon'\Lambda^{-1}\varepsilon\right) + \frac{1}{|\Lambda|^{1/2}}[\Lambda^{-1}\varepsilon\varepsilon'\Lambda^{-1}]_{rt}\exp\left(-\frac{1}{2}\varepsilon'\Lambda^{-1}\varepsilon\right)\right\}$$
$$= f(\varepsilon \mid \Lambda)\{[\Lambda^{-1}\varepsilon\varepsilon'\Lambda^{-1}]_{rt} - [\Lambda^{-1}]_{rt}\}. \quad \text{(A9)}$$

Define the following notations:

$W_{is} = (Z_i'x_{s1}, Z_i'x_{s2}, \ldots, Z_i'x_{sJ})'$,

$\varphi_{isj} = f(y_{is} = j \mid \theta, \Lambda) = \int p_{isj}(\varepsilon) f(\varepsilon \mid \Lambda)\, d\varepsilon$,

$p_{is} = (p_{is1}(\varepsilon), \ldots, p_{isJ}(\varepsilon))'$,

$P_{is}(\varepsilon) = \mathrm{diag}(p_{is1}(\varepsilon), \ldots, p_{isJ}(\varepsilon))$,

$\varsigma(\varepsilon) =$ a vector of length $m(m+1)/2$ with elements $\{\frac{1}{2}[\Lambda^{-1}\varepsilon\varepsilon'\Lambda^{-1}]_{rr} - \frac{1}{2}[\Lambda^{-1}]_{rr}\}$ and $\{[\Lambda^{-1}\varepsilon\varepsilon'\Lambda^{-1}]_{rt} - [\Lambda^{-1}]_{rt}\}$, for $r < t$, $r, t \in \{1, \ldots, m\}$.

Using these notations and working backward starting from Equations (A9) to (A1), we get

$$I(\theta, \lambda_B) = \sum_{i=1}^{N}\sum_{s=1}^{S}\begin{bmatrix} A_{is}'\Delta_{is}^{-1}A_{is} & A_{is}'\Delta_{is}^{-1}\Gamma_{is} \\ \Gamma_{is}'\Delta_{is}^{-1}A_{is} & \Gamma_{is}'\Delta_{is}^{-1}\Gamma_{is}\end{bmatrix}, \quad \text{where}$$

$$\Delta_{is} = \mathrm{diag}(\varphi_{is1}, \ldots, \varphi_{isJ}),$$

$$A_{is} = \int [\{P_{is}(\varepsilon) - p_{is}(\varepsilon)p_{is}(\varepsilon)'\}]W_{is} f(\varepsilon \mid \Lambda)\, d\varepsilon,$$

$$\Gamma_{is} = \int p_{is}(\varepsilon)\varsigma(\varepsilon)' f(\varepsilon \mid \Lambda)\, d\varepsilon.$$

## Appendix B. Information Matrix for the Consider-Then-Choose Model

In this appendix, we show the derivation of the information matrix $I(\theta, \lambda_B, \zeta)$ for the consider-then-choose model where $\lambda_B$ denotes the vector of the $m(m+1)/2$ unique parameters in $\Lambda$, and $\zeta = \{\xi_1, \ldots \xi_R\}$ is the vector that consists of the $R$ screening probabilities.

The Fisher information matrix $I(\theta, \lambda_B, \zeta)$ can be expressed as

$$I(\theta, \lambda_B, \zeta) = \begin{bmatrix} -E\frac{\partial^2 \log L(\theta,\Lambda,\zeta)}{\partial\theta\,\partial\theta'} & -E\frac{\partial^2 \log L(\theta,\Lambda,\zeta)}{\partial\theta\,\partial\lambda_B'} & -E\frac{\partial^2 \log L(\theta,\Lambda,\zeta)}{\partial\theta\,\partial\zeta'} \\ -E\frac{\partial^2 \log L(\theta,\Lambda,\zeta)}{\partial\lambda_B\,\partial\theta'} & -E\frac{\partial^2 \log L(\theta,\Lambda,\zeta)}{\partial\lambda_B\,\partial\lambda_B'} & -E\frac{\partial^2 \log L(\theta,\Lambda,\zeta)}{\partial\lambda_B\,\partial\zeta'} \\ -E\frac{\partial^2 \log L(\theta,\Lambda,\zeta)}{\partial\zeta\,\partial\theta'} & -E\frac{\partial^2 \log L(\theta,\Lambda,\zeta)}{\partial\zeta\,\partial\lambda_B'} & -E\frac{\partial^2 \log L(\theta,\Lambda,\zeta)}{\partial\zeta\,\partial\zeta'}\end{bmatrix}. \quad \text{(B1)}$$

Following the derivation of (A2)–(A4), it can be shown that, similarly,

$$E\frac{\partial^2}{\partial\theta\partial\theta'}\log L(\theta,\Lambda,\zeta) = -\sum_{i=1}^{N}\sum_{s=1}^{S}\sum_{j=0}^{J}\frac{(\partial f(y_{is}=j\mid\theta,\Lambda,\zeta)/\partial\theta)(\partial f(y_{is}=j\mid\theta,\Lambda,\zeta)/\partial\theta')}{f(y_{is}=j\mid\theta,\Lambda,\zeta)}, \quad \text{(B2)}$$

$$E\frac{\partial^2}{\partial\theta\partial\lambda_B'}\log L(\theta,\Lambda,\zeta) = -\sum_{i=1}^{N}\sum_{s=1}^{S}\sum_{j=0}^{J}\frac{(\partial f(y_{is}=j\mid\theta,\Lambda,\zeta)/\partial\theta)(\partial f(y_{is}=j\mid\theta,\Lambda,\zeta)/\partial\lambda_B')}{f(y_{is}=j\mid\theta,\Lambda,\zeta)}, \quad \text{(B3)}$$

$$E\frac{\partial^2}{\partial\lambda_B\partial\lambda_B'}\log L(\theta,\Lambda,\zeta) = -\sum_{i=1}^{N}\sum_{s=1}^{S}\sum_{j=0}^{J}\frac{(\partial f(y_{is}=j\mid\theta,\Lambda,\zeta)/\partial\lambda_B)(\partial f(y_{is}=j\mid\theta,\Lambda,\zeta)/\partial\lambda_B')}{f(y_{is}=j\mid\theta,\Lambda,\zeta)}, \quad \text{(B4)}$$

$$E\frac{\partial^2}{\partial\theta\partial\zeta'}\log L(\theta,\Lambda,\zeta) = -\sum_{i=1}^{N}\sum_{s=1}^{S}\sum_{j=0}^{J}\frac{(\partial f(y_{is}=j\mid\theta,\Lambda,\zeta)/\partial\theta)(\partial f(y_{is}=j\mid\theta,\Lambda,\zeta)/\partial\zeta')}{f(y_{is}=j\mid\theta,\Lambda,\zeta)}, \quad \text{(B5)}$$

$$E\frac{\partial^2}{\partial\lambda_B\partial\zeta'}\log L(\theta,\Lambda,\zeta) = -\sum_{i=1}^{N}\sum_{s=1}^{S}\sum_{j=0}^{J}\frac{(\partial f(y_{is}=j\mid\theta,\Lambda,\zeta)/\partial\lambda_B)(\partial f(y_{is}=j\mid\theta,\Lambda,\zeta)/\partial\zeta')}{f(y_{is}=j\mid\theta,\Lambda,\zeta)}, \quad \text{(B6)}$$

$$E\frac{\partial^2}{\partial\zeta\partial\zeta'}\log L(\theta,\Lambda,\zeta) = -\sum_{i=1}^{N}\sum_{s=1}^{S}\sum_{j=0}^{J}\frac{(\partial f(y_{is}=j\mid\theta,\Lambda,\zeta)/\partial\zeta)(\partial/\partial f(y_{is}=j\mid\theta,\Lambda,\zeta)\zeta')}{f(y_{is}=j\mid\theta,\Lambda,\zeta)}. \quad \text{(B7)}$$

Note that $j = 0$ represents the no-choice option, and

$$f(y_{is}=j\mid\theta,\Lambda,\zeta) = \int\sum_{\tau=\tau^1}^{\tau^{2^R}}\left\{p_{isj}(\boldsymbol{\tau},\varepsilon)\prod_{r=1}^{R}\xi_r^{I\{\tau_r=1\}}(1-\xi_r)^{I\{\tau_r=0\}}\right\}f(\varepsilon\mid\Lambda)\,d\varepsilon, \quad \text{(B8)}$$

$$\frac{\partial}{\partial\theta}f(y_{is}=j\mid\theta,\Lambda,\zeta) = \int\sum_{\tau=\tau^1}^{\tau^{2^R}}\left\{\Theta_{isj}(\boldsymbol{\tau},\varepsilon)\prod_{r=1}^{R}\xi_r^{I\{\tau_r=1\}}(1-\xi_r)^{I\{\tau_r=0\}}\right\}f(\varepsilon\mid\Lambda)\,d\varepsilon, \quad \text{(B9)}$$

where $\Theta_{isj}(\boldsymbol{\tau}, \varepsilon) = 0$ if the consideration set is empty, i.e., if $C_{is}(\boldsymbol{\tau}) = \varnothing$; otherwise, $\Theta_{isj}(\boldsymbol{\tau}, \varepsilon) = p_{isj}(\boldsymbol{\tau},\varepsilon)Z_i'x_{sj} - p_{isj}(\boldsymbol{\tau},\varepsilon)\cdot\sum_{k\in C_{is}(\tau)}p_{isk}(\boldsymbol{\tau},\varepsilon)Z_i'x_{sk}$.

$$\frac{\partial}{\partial\lambda_B}f(y_{is}=j\mid\theta,\Lambda,\zeta) = \int\sum_{\tau=\tau^1}^{\tau^{2^R}}\left\{p_{isj}(\tau,\varepsilon)\prod_{r=1}^{R}\xi_r^{I\{\tau_r=1\}}(1-\xi_r)^{I\{\tau_r=0\}}\right\}\frac{\partial f(\varepsilon\mid\Lambda)d\varepsilon}{\partial\lambda_B}, \quad \text{(B10)}$$

where $(\partial f(\varepsilon \mid \Lambda)/\partial\lambda_B)$ is the same as shown in (A8) and (A9) of Appendix A for the different elements of vector $\lambda_B$.

For each element in vector $\zeta = \{\xi_1, \ldots, \xi_R\}$, we have

$$\frac{\partial}{\partial\xi_k}f(y_i=j\mid\theta,\Lambda,\zeta) = \int\sum_{\tau=\tau^1}^{\tau^{2^R}}p_{isj}(\boldsymbol{\tau},\varepsilon)\eta_k(\boldsymbol{\tau})f(\varepsilon\mid\Lambda)\,d\varepsilon, \quad k = 1, \ldots, R, \quad \text{(B11)}$$

where

$$\eta_k(\tau) = \Big\{ I\{\tau_k = 1\} \prod_{r\neq k} \xi_r^{I\{\tau_r=1\}}(1-\xi_r)^{I\{\tau_r=0\}} - I\{\tau_k = 0\} \prod_{r\neq k} \xi_r^{I\{\tau_r=1\}}(1-\xi_r)^{I\{\tau_r=0\}} \Big\}.$$

Define the following notations:

$$W_{is} = (0, Z_i' x_{s1}, Z_i' x_{s2}, \ldots, Z_i' x_{sJ})',$$

$$\eta(\boldsymbol{\tau}) = (\eta_1(\boldsymbol{\tau}), \ldots, \eta_R(\boldsymbol{\tau}))',$$

$$P_{is}(\boldsymbol{\tau}, \varepsilon) = \operatorname{diag}(p_{is0}(\boldsymbol{\tau}, \varepsilon), \ldots, p_{isj}(\boldsymbol{\tau}, \varepsilon)),$$

$$p_{is}(\boldsymbol{\tau}, \varepsilon) = (p_{is0}(\boldsymbol{\tau}, \varepsilon), \ldots, p_{isj}(\boldsymbol{\tau}, \varepsilon))',$$

$$\pi_{isj} = f(y_{is} = j \mid \theta, \Lambda, \zeta) = \int \sum_{\tau=\tau^1}^{\tau^{2^R}} \Big\{ p_{isj}(\boldsymbol{\tau}, \varepsilon) \prod_{r=1}^{R} \xi_r^{I\{\tau_r=1\}}(1-\xi_r)^{I\{\tau_r=0\}} \Big\} f(\varepsilon \mid \Lambda)\, d\varepsilon.$$

$\varsigma(\varepsilon)$ = a vector of length $m(m+1)/2$ with elements $\{\frac{1}{2}[\Lambda^{-1}\varepsilon\varepsilon'\Lambda^{-1}]_{rr} - \frac{1}{2}[\Lambda^{-1}]_{rr}\}$ and $\{[\Lambda^{-1}\varepsilon\varepsilon'\Lambda^{-1}]_{rt} - [\Lambda^{-1}]_{rt}\}$, for $r < t$, $r, t \in \{1, \ldots, m\}$.

Using these notations and working backward starting from (B11) to (B1), we get

$$I(\theta, \lambda_B, \zeta) = \sum_{i=1}^{N} \sum_{s=1}^{S} \begin{pmatrix} M_{is}'\Pi_{is}^{-1}M_{is} & M_{is}'\Pi_{is}^{-1}\Omega_{is} & M_{is}'\Pi_{is}^{-1}\Xi_{is} \\ \Omega_{is}'\Pi_{is}^{-1}M_{is} & \Omega_{is}'\Pi_{is}^{-1}\Omega_{is} & \Omega_{is}'\Pi_{is}^{-1}\Xi_{is} \\ \Xi_{is}'\Pi_{is}^{-1}M_{is} & \Xi_{is}'\Pi_{is}^{-1}\Omega_{is} & \Xi_{is}'\Pi_{is}^{-1}\Xi_{is} \end{pmatrix},$$

where

$$M_{is} = \int \sum_{\tau=\tau^1}^{\tau^{2\wedge R}} \Big\{ [P_{is}(\boldsymbol{\tau}, \varepsilon) - p_{is}(\boldsymbol{\tau}, \varepsilon) p_{is}(\boldsymbol{\tau}, \varepsilon)'] \cdot \prod_{r=1}^{R} \xi_r^{I\{\tau_r=1\}}(1-\xi_r)^{I\{\tau_r=0\}} \Big\} W_{is} f(\varepsilon \mid \Lambda)\, d\varepsilon,$$

$$\Omega_{is} = \int \sum_{\tau=\tau^1}^{\tau^{2\wedge R}} \Big\{ p_{is}(\boldsymbol{\tau}, \varepsilon) \prod_{r=1}^{R} \xi_r^{I\{\tau_r=1\}}(1-\xi_r)^{I\{\tau_r=0\}} \Big\} \varsigma(\varepsilon)' f(\varepsilon \mid \Lambda)\, d\varepsilon,$$

$$\Xi_{is} = \int \sum_{\tau=\tau^1}^{\tau^{2\wedge R}} \Big\{ p_{is}(\boldsymbol{\tau}, \varepsilon) \eta(\boldsymbol{\tau})' \Big\} f(\varepsilon \mid \Lambda)\, d\varepsilon,$$

$$\Pi_{is} = \operatorname{diag}(\pi_{is0}, \ldots, \pi_{isj}).$$

## Appendix C. Design Examples for the Low Preference Scenario

| | | Compensatory design overlap rate = 69.8% | | | | Conjunctive design, low screening overlap rate = 71.9% | | | | Conjunctive design, high screening overlap rate = 86.5% | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Attributes | | | | Attributes | | | | Attributes | | | |
| Choice set | Profile | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| 1 | I | 2 | 3 | 2 | 3 | 1 | 1 | 1 | 2 | 2 | 1 | 1 | 3 |
| | II | 2 | 3 | 1 | 3 | 3 | 1 | 2 | 2 | 1 | 1 | 1 | 1 |
| | III | 3 | 3 | 1 | 1 | 3 | 3 | 2 | 2 | 3 | 1 | 1 | 1 |
| 2 | I | 2 | 1 | 2 | 2 | 2 | 3 | 2 | 3 | 1 | 3 | 1 | 1 |
| | II | 1 | 1 | 3 | 3 | 3 | 3 | 1 | 3 | 1 | 2 | 2 | 1 |
| | III | 3 | 2 | 3 | 3 | 2 | 1 | 3 | 1 | 2 | 1 | 3 | 1 |
| 3 | I | 2 | 2 | 1 | 3 | 1 | 1 | 1 | 1 | 1 | 3 | 2 | 2 |
| | II | 1 | 3 | 2 | 1 | 1 | 3 | 3 | 3 | 1 | 2 | 2 | 1 |
| | III | 2 | 3 | 3 | 1 | 3 | 1 | 1 | 2 | 1 | 1 | 2 | 1 |
| 4 | I | 1 | 2 | 2 | 2 | 1 | 1 | 3 | 1 | 1 | 2 | 2 | 1 |
| | II | 1 | 3 | 3 | 1 | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 3 |
| | III | 2 | 3 | 3 | 3 | 1 | 2 | 2 | 3 | 2 | 2 | 3 | 2 |
| 5 | I | 2 | 2 | 3 | 1 | 1 | 1 | 3 | 2 | 2 | 1 | 2 | 1 |
| | II | 2 | 1 | 3 | 2 | 1 | 3 | 2 | 2 | 2 | 1 | 1 | 3 |
| | III | 1 | 3 | 1 | 1 | 1 | 3 | 1 | 1 | 2 | 3 | 2 | 3 |
| 6 | I | 3 | 1 | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 3 | 1 |
| | II | 1 | 2 | 3 | 1 | 1 | 1 | 1 | 3 | 2 | 3 | 3 | 1 |
| | III | 3 | 1 | 1 | 3 | 3 | 3 | 3 | 1 | 3 | 2 | 3 | 1 |
| 7 | I | 1 | 2 | 1 | 2 | 2 | 2 | 1 | 2 | 1 | 3 | 3 | 1 |
| | II | 3 | 3 | 2 | 1 | 3 | 2 | 2 | 1 | 2 | 2 | 1 | 3 |
| | III | 1 | 1 | 3 | 2 | 1 | 2 | 3 | 3 | 1 | 2 | 1 | 1 |
| 8 | I | 3 | 1 | 2 | 3 | 2 | 1 | 2 | 1 | 2 | 1 | 1 | 3 |
| | II | 2 | 1 | 2 | 1 | 3 | 2 | 1 | 1 | 1 | 3 | 1 | 3 |
| | III | 2 | 3 | 1 | 2 | 1 | 3 | 3 | 1 | 2 | 3 | 3 | 3 |
| 9 | I | 3 | 3 | 2 | 1 | 3 | 2 | 3 | 2 | 3 | 2 | 1 | 3 |
| | II | 2 | 1 | 1 | 1 | 3 | 2 | 2 | 3 | 1 | 2 | 1 | 3 |
| | III | 1 | 2 | 3 | 1 | 3 | 3 | 2 | 1 | 3 | 1 | 3 | 1 |

**Appendix C. (Cont'd.)**

| | | Compensatory design overlap rate = 69.8% | | | | Conjunctive design, low screening overlap rate = 71.9% | | | | Conjunctive design, high screening overlap rate = 86.5% | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Attributes | | | | Attributes | | | | Attributes | | | |
| Choice set | Profile | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| 10 | I | 2 | 2 | 3 | 2 | 2 | 1 | 1 | 2 | 1 | 1 | 1 | 1 |
| | II | 2 | 2 | 2 | 3 | 2 | 2 | 1 | 1 | 1 | 1 | 3 | 2 |
| | III | 3 | 3 | 3 | 2 | 2 | 1 | 2 | 3 | 1 | 2 | 3 | 1 |
| 11 | I | 3 | 2 | 3 | 1 | 2 | 1 | 1 | 1 | 1 | 1 | 2 | 1 |
| | II | 3 | 1 | 2 | 1 | 3 | 1 | 2 | 1 | 1 | 3 | 1 | 1 |
| | III | 2 | 3 | 3 | 2 | 1 | 1 | 2 | 1 | 3 | 1 | 2 | 3 |
| 12 | I | 3 | 1 | 1 | 2 | 2 | 2 | 1 | 1 | 2 | 3 | 1 | 2 |
| | II | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 1 | 3 | 3 | 3 | 1 |
| | III | 1 | 3 | 1 | 3 | 3 | 2 | 3 | 2 | 3 | 3 | 2 | 2 |
| 13 | I | 2 | 3 | 2 | 3 | 1 | 3 | 1 | 3 | 3 | 1 | 1 | 2 |
| | II | 1 | 1 | 3 | 2 | 2 | 1 | 3 | 3 | 3 | 1 | 2 | 2 |
| | III | 1 | 1 | 2 | 3 | 1 | 3 | 1 | 2 | 3 | 3 | 2 | 1 |
| 14 | I | 2 | 2 | 3 | 1 | 2 | 3 | 1 | 3 | 1 | 2 | 2 | 2 |
| | II | 2 | 3 | 1 | 3 | 1 | 1 | 2 | 3 | 1 | 2 | 3 | 2 |
| | III | 3 | 2 | 3 | 2 | 1 | 2 | 3 | 3 | 2 | 1 | 2 | 1 |
| 15 | I | 3 | 2 | 2 | 2 | 2 | 2 | 2 | 3 | 2 | 2 | 2 | 1 |
| | II | 1 | 1 | 3 | 1 | 1 | 3 | 2 | 2 | 1 | 3 | 3 | 3 |
| | III | 1 | 3 | 2 | 1 | 1 | 2 | 2 | 1 | 3 | 3 | 2 | 1 |
| 16 | I | 1 | 1 | 1 | 2 | 3 | 2 | 1 | 2 | 1 | 2 | 1 | 2 |
| | II | 3 | 1 | 3 | 1 | 3 | 3 | 3 | 3 | 3 | 2 | 2 | 3 |
| | III | 1 | 3 | 2 | 3 | 3 | 2 | 2 | 3 | 1 | 2 | 3 | 2 |
| 17 | I | 1 | 1 | 1 | 1 | 1 | 3 | 1 | 1 | 3 | 1 | 1 | 2 |
| | II | 3 | 2 | 3 | 2 | 2 | 1 | 1 | 3 | 1 | 1 | 2 | 2 |
| | III | 2 | 1 | 2 | 3 | 2 | 2 | 1 | 3 | 1 | 1 | 1 | 3 |
| 18 | I | 1 | 3 | 3 | 2 | 1 | 2 | 2 | 1 | 2 | 2 | 1 | 1 |
| | II | 3 | 1 | 3 | 3 | 3 | 2 | 1 | 1 | 2 | 2 | 2 | 2 |
| | III | 3 | 3 | 1 | 3 | 2 | 3 | 2 | 2 | 3 | 2 | 3 | 3 |
| 19 | I | 3 | 2 | 2 | 3 | 3 | 1 | 3 | 1 | 1 | 1 | 1 | 2 |
| | II | 3 | 1 | 1 | 1 | 1 | 3 | 1 | 2 | 1 | 2 | 1 | 2 |
| | III | 3 | 3 | 3 | 2 | 1 | 2 | 3 | 2 | 2 | 3 | 2 | 1 |
| 20 | I | 2 | 2 | 2 | 1 | 2 | 3 | 1 | 2 | 2 | 3 | 2 | 2 |
| | II | 2 | 1 | 2 | 1 | 3 | 2 | 1 | 2 | 1 | 1 | 2 | 2 |
| | III | 3 | 3 | 3 | 3 | 3 | 1 | 1 | 3 | 3 | 1 | 1 | 1 |
| 21 | I | 1 | 3 | 1 | 3 | 1 | 2 | 3 | 2 | 3 | 1 | 1 | 1 |
| | II | 2 | 2 | 3 | 2 | 3 | 1 | 3 | 2 | 3 | 3 | 1 | 1 |
| | III | 3 | 2 | 1 | 1 | 2 | 1 | 3 | 3 | 2 | 1 | 1 | 1 |
| 22 | I | 3 | 1 | 3 | 1 | 1 | 3 | 3 | 1 | 2 | 2 | 1 | 2 |
| | II | 2 | 3 | 1 | 1 | 1 | 3 | 1 | 3 | 1 | 2 | 2 | 2 |
| | III | 1 | 1 | 3 | 3 | 1 | 2 | 1 | 1 | 2 | 1 | 3 | 3 |
| 23 | I | 2 | 3 | 1 | 3 | 2 | 2 | 2 | 3 | 2 | 1 | 2 | 2 |
| | II | 1 | 2 | 2 | 3 | 3 | 2 | 2 | 3 | 2 | 1 | 1 | 1 |
| | III | 1 | 1 | 2 | 1 | 2 | 3 | 3 | 1 | 2 | 2 | 1 | 2 |
| 24 | I | 1 | 1 | 2 | 2 | 1 | 3 | 2 | 2 | 1 | 2 | 1 | 2 |
| | II | 3 | 3 | 3 | 1 | 2 | 3 | 3 | 3 | 1 | 3 | 1 | 3 |
| | III | 1 | 3 | 1 | 1 | 3 | 3 | 2 | 1 | 1 | 2 | 3 | 3 |

## References

Allenby, G. M., N. Arora, J. L. Ginter. 1998. On the heterogeneity of demand. *J. Marketing Res.* **35**(3) 384–389.

Araña, J. E., C. J. León, M. W. Hanemann 2008. Emotions and decision rules in discrete choice experiments for valuing health care programmes for the elderly. *J. Health Econom.* **27**(3) 753–769.

Arora, N., J. Huber. 2001. Improving parameter estimates and model prediction by aggregate customization in choice experiments. *J. Consumer Res.* **28**(2) 273–283.

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

Bradlow, E. T., V. R. Rao. 2000. A hierarchical Bayes model for assortment choice. *J. Marketing Res.* **37**(2) 259–268.

Gilbride, T. J., G. M. Allenby. 2004. A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Sci.* **23**(3) 391–406.

Gilbride, T. J., G. M. Allenby. 2006. Estimating heterogeneous EBA and economic screening rule choice models. *Marketing Sci.* **25**(5) 494–509.

Green, P. E. 1974. On the design of choice expedments involving multifactor alternatives. *J. Consumer Res.* **1**(2) 61–68.

Hauser, J. R., O. Toubia, T. Evgeniou, R. Befurt, D. Dzyabura. 2010. Disjunctions of conjunctions, cognitive simplicity and consideration sets. *J. Marketing Res.* **47**(3) 485–496.

Hess, S., K. E. Train, J. W. Polak. 2006. On the use of a modified Latin hypercube sampling (MLHS) method in the estimation of a mixed logit model for vehicle choice. *Transportation Res. Part B* **40**(2) 147–163.

Huber, J., K. Zwerina. 1996. The importance of utility balance in efficient choice designs. *J. Marketing Res.* **33**(3) 307–317.

Jedidi, K. R. Kohli. 2005. Probabilistic subset-conjunctive models for heterogeneous consumers. *J. Marketing Res.* **42**(4) 483–495.

Jedidi, K., R. Kohli, W. S. DeSarbo. 1996. Consideration sets in conjoint analysis. *J. Marketing Res.* **33**(3) 364–372.

Johnson, R. 2008. A perspective on adaptive CBC (what can we expect from respondents?). Sawtooth Software technical paper, Sequium, WA. http://www.sawtoothsoftware.com/download/techpap/acbcperspective.pdf.

Kessels, R., P. Goos, M. Vandebroek. 2006. A comparison of criteria to design efficient choice experiments. *J. Marketing Res.* **43**(3) 409–419.

Kessels, R., B. Jones, P. Goos, M. Vanderbroek. 2009. An efficient algorithm for constructing Bayesian optimal choice designs. *J. Bus. Econom. Statist.* **27**(2) 279–291.

Kohli, R., K. Jedidi. 2005. Probabilistic subset conjunction. *Psychometrika* **70**(4) 737–757.

Kohli, R., K. Jedidi. 2007. Representation and inference of lexicographic preference models and their variants. *Marketing Sci.* **26**(3) 380–399.

Kuhfeld, W. F., R. Tobias. 2005. Large factorial designs for product engineering and marketing research applications. *Technometrics* **47**(2) 132–141.

Lussier, D. A., R. W. Olshavsky. 1979. Task complexity and contingent processing in brand choice. *J. Consumer Res.* **6**(2) 154–165.

McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. P. Zarembka, ed. *Frontiers in Econometrics*. Academic Press, New York, 105–142.

Meyer, R. K., C. J. Nachtsheim. 1995. The coordinate-exchange algorithm for constructing exact optimal experimental designs. *Technometrics* **37**(1) 60–69.

Newton, M., A. Raftery. 1994. Approximate Bayesian inference with weighted likelihood bootstrap. *J. Roy. Statist. Soc. Ser. B* **56**(1) 3–48.

Orme, B. 2009. Fine-tuning CBC and adaptive CBC questionnaires. Sawtooth Software technical paper, Sequium, WA. http://www.sawtooth-software.com/download/techpap/finetune.pdf.

Payne, J. L. 1976. Task complexity and contingent processing in decision making—An information search and protocol analysis. *Organ. Behav. Human Performance* **16**(2) 366–387.

Rossi, P. E., G. M. Allenby, R. McCulloch. 2005. *Bayesian Statistics and Marketing*. John Wiley & Sons, Chichester, UK.

Sándor, Z., M. Wedel. 2001. Designing conjoint choice experiments using managers' prior beliefs. *J. Marketing Res.* **38**(4) 430–444.

Sándor, Z., M. Wedel. 2002. Profile construction in experimental choice designs for mixed logit models. *Marketing Sci.* **21**(4) 455–475.

Sándor, Z., M. Wedel. 2005. Heterogeneous conjoint choice designs. *J. Marketing Res.* **42**(2) 210–218.

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, A. van der Linde. 2002. Bayesian measures of model complexity and fit. *J. Roy. Statist. Soc. Ser. B* **64**(4) 583–639.

Sun, F., M.-Q. Liu, D. K. J. Lin. 2009. Construction of orthogonal Latin hypercube designs. *Biometrika* **96**(4) 971–974.

Swait, J. 2001. A non-compensatory choice model incorporating attribute cutoffs. *Transportation Res. Part B* **35**(10) 903–928.

Toubia, O., J. R. Hauser. 2007. On managerially efficient experimental designs. *Marketing Sci.* **26**(6) 851–858.

van Nierop, E., B. Bronnenberg, R. Paap, M. Wedel, P. H. Franses. 2010. Retrieving unobserved consideration sets from household panel data. *J. Marketing Res.* **47**(1) 63–74.

Wedel, M., W. A. Kamakura, N. Arora, A. Bemmaor, J. Chiangm, T. Elrod, R. Johnson, P. Lenk, S. Neslin, C. S. Poulsen. 1999. Discrete and continuous representation of heterogeneity. *Marketing Lett.* **10**(3) 217–230.

Yu, J., P. Goos, M. Vanderbroek. 2009. Efficient conjoint choice designs in the presence of respondent heterogeneity. *Marketing Sci.* **28**(1) 122–135.

Zhu, W., H. Timmermans. 2009. Modelling pedestrian go-home decisions: A comparison of linear and nonlinear compensatory, and conjunctive non-compensatory specifications. *J. Retailing Consumer Services* **16**(3) 227–231.