TY HENDERSON and QING LIU*

Variable selection is a decision heuristic that describes a selective choice process in which choices are made on the basis of only a subset of product attributes while the presence of other ("inactive") attributes plays no active role in the decision. Within this context, the authors address two integrated topics that have received scant attention: the efficient design of choice experiments and the analysis of data that arises from a selective choice process. The authors propose a new dual-objective compound design criterion that incorporates prior information for the joint purpose of efficiently estimating the effects of the active attributes and detecting the effects of attributes labeled as inactive that may turn out to be active. The approach leverages self-stated auxiliary data as prior information both for individual-level customized design construction and in a heterogeneous variable selection model. The authors demonstrate the efficiency advantages of the approach relative to design benchmarks and highlight practical implications using both simulated data and actual data from a conjoint choice experiment in which individual designs were customized instantaneously using self-stated active–inactive attribute status.

Keywords: variable selection, selective choice process, customized conjoint choice designs, compound design criterion, heterogeneous variable selection model

Online Supplement: http://dx.doi.org/10.1509/jmr.13.0545

Efficient Design and Analysis for a Selective Choice Process

People often make choices using only a subset of product attributes; that is, there is variable selection in the choice decision process (Gilbride, Allenby, and Brazell 2006; Hensher, Rose, and Greene 2012). Inactive (unselected) attributes do not contribute to the expected utility for a choice alternative and could be described as attributes that "do not matter." In the marketing literature, models of choice process decision heuristics (e.g., screening rules [Gilbride and Allenby 2004], consideration sets [Van Nierop et al. 2010], lexicographic search [Yee et al. 2007]) have advanced understanding of how people simplify the choice process subject to cognitive constraints (for details, see Bettman, Luce, and Payne's [1998] review), but less attention has been paid to the efficient design of choice experiments under such decision heuristics (Liu and Arora 2011). Within the variable selection literature (e.g., George and McCulloch 1993; Tibshirani 1996; Zou and Hastie 2005), this article addresses two integrated topics that have received scant attention: the efficient design of choice experiments and the analysis of data from a selective choice process.

Data from a selective choice process present unique challenges. In contrast to an analysis focused only on preferences as measured by attribute-level partworths (i.e., β), the goal of design for a selective choice process across a heterogeneous population is twofold: it aims for (1) the active–inactive classification of all attributes and (2) the efficient estimation of the partworths of the active attributes. In the spirit of combining data in hybrid models (Green and Krieger 1996), our proposed solution leverages self-stated auxiliary data as prior information in both individual-level customized design construction and a heterogeneous variable selection model.

This article introduces the compound design criterion (Atkinson, Donev, and Tobias 2007) to the marketing literature. In contrast to extant research on efficient choice designs

*Ty Henderson is Associate Professor of Marketing, McCombs School of Business, University of Texas at Austin (e-mail: ty.henderson@mccombs.utexas.edu). Qing Liu is Associate Professor of Marketing, Wisconsin School of Business, University of Wisconsin–Madison (e-mail: qliu@bus.wisc.edu). Coeditor: Robert Meyer; Associate Editor: Eric Bradlow.

ISSN: 0022-2437 (print)
1547-7193 (electronic)

Journal of Marketing Research
Vol. LIV (June 2017), 430–446
DOI: 10.1509/jmr.13.0545

that focus on a single objective—the efficient estimation of all model parameters (e.g., the D-criterion in Arora and Huber [2001], Huber and Zwerina [1996], and Sándor and Wedel [2005]), we propose a compound design criterion that balances two design objectives: efficient estimation of the effects of the active attributes and detection of the effects of the attributes that are labeled inactive but may turn out to be active. We customize a respondent-level experimental design on the basis of self-stated information as to which attributes are likely to be active in the respondent's selective choice process.

The results show that designs obtained using the proposed design criterion are more efficient than a set of comparable D-criterion designs. Our analysis of design characteristic differences provides intuition as to why the proposed criterion offers advantages in design efficiency. Because design construction using the proposed criterion requires a priori assumptions regarding active–inactive attribute status, we also investigate design robustness and find that the efficiency advantages of the proposed designs over the benchmark designs generally hold even when such assumptions are moderately misspecified. A simulation study shows that when the choice process is selective, the proposed approach more faithfully recovers individual preferences and provides better market predictions. Results from a conjoint choice experiment demonstrate the feasibility of a customization strategy using the proposed design criterion.

The business case for analyzing a selective choice process is often made in high-dimensional contexts in which people sort through numerous attributes, some of which may not matter to everyone. When assessing market readiness for a product using conjoint choice experiments, failure to account for the possibility of a selective choice process could distort inference on these potentially active–inactive attributes, particularly when there is heterogeneity in attribute status across people in the market (Hensher and Rose 2009). Our approach to design and analysis for a selective choice process incorporates respondent-specific information about the active–inactive status of an attribute, customizes a design to match the presumed active–inactive pattern, and more faithfully recovers heterogeneous preference partworths near zero for inactive attributes or away from zero for active attributes.

The remainder of the article is organized as follows. In the next section, we present the proposed compound design criterion for the construction of efficient designs for a selective choice process and compare the proposed design criterion with benchmark D-criterion designs under a range of variable selection contexts. We then describe the heterogeneous variable selection model that incorporates auxiliary data, explain our approach to design customization that utilizes self-stated information from each respondent, and use a simulation to demonstrate the superiority of the new design approach. We follow this with an empirical application showing the execution of the customized design strategy and close with a discussion of cost–benefit considerations for design and analysis implementation.

EFFICIENT DESIGNS FOR A SELECTIVE CHOICE PROCESS

The Proposed Design Criterion

We propose a compound design criterion (Atkinson, Donev, and Tobias 2007, Chapter 21) that customizes the experimental design for each respondent on the basis of self-stated prior information regarding which attributes are likely to be active in his or her selective choice process. Specifically, assume that for a given person, the expected consumer utility is

$$V = X\beta = X_1\beta_1 + X_2\beta_2, \tag{1}$$

where X_1 represents the set of variables corresponding to the attributes that are likely to be active during the respondent's decision process, and X_2 represents the set of variables corresponding to the attributes that are likely to be inactive based on prior information from the respondent (e.g., the respondent's self-stated information on attributes that "matter" or "do not matter"). The total number of variables (or the size of the parameter vector β) is $m = m_1 + m_2$, where m_1 is the size of the parameter vector β_1, and m_2 is the size of the parameter vector β_2. Note that if the parameter vector β_2 is equal to 0, then the full model ($X\beta$) specification of utility is equivalent to the reduced model ($X_1\beta_1$) that includes only the set of active variables. Let $I(\beta|X)$ represent the Fisher information matrix obtained under the full model (e.g., Sándor and Wedel 2001), that is,

$$I(\beta|X) = \sum_{s=1}^{S} X_s'\left[P_s - p_s p_s'\right]X_s, \tag{2}$$

where S is the total number of choice sets, J is the number of alternatives per choice set, and $X_s = (x_{s1}, x_{s2}, \ldots, x_{sJ})'$, $P_s(v) = \mathrm{diag}(p_{s1}, \ldots, p_{sJ})$, $p_s = (p_{s1}, \ldots, p_{sJ})'$, with choice probabilities $p_{sj} = \exp(x_{sj}'\beta)/\sum_{j=1}^{J}\exp(x_{sj}'\beta)$, $j = 1, \ldots, J$. Similarly, let $I(\beta_1|X_1)$ represent the Fisher information matrix obtained under the reduced model. In addition, let $I(\beta_1|X)$ represent the first $m_1 \times m_1$ submatrix of the Fisher information matrix $I(\beta|X)$ obtained under the full model. We define the following compound Bayesian design criterion, which we label as the ϕ-criterion, to construct designs for the joint purpose of efficient estimation of the effects of active variables and detection of the effects of variables that are stated as inactive but may turn out to be active:

$$\phi = \int\{w[\log|I(\beta_1|X_1)|] + (1-w)[\log|I(\beta|X)| - \log|I(\beta_1|X)|]\}\,\pi(\beta)d\beta. \tag{3}$$

The weight w in Equation 3 is to be specified by the experimenter. It indicates the relative emphasis that is placed on the two objectives in the compound design criterion. For example, if w = 1, then all emphasis is placed on the efficient estimation of the effects of active variables in the reduced model only. If, instead, w = .8, then 80% of the emphasis is placed on the efficient estimation of the effects of active variables in the reduced model (the first component of the ϕ-criterion), while 20% of the emphasis is on the detection of the effects of those attributes in the full model that are stated to be inactive but may turn out to be active (the second component of the ϕ-criterion). The distinction between the active–inactive elements contained in X_1 and X_2 is defined in this setup by self-stated auxiliary data from individual respondents, which gives rise to designs that are customized for respondents.

Because the value of the Fisher information matrix $I(\beta|X)$, as shown in Equation 3, is dependent on the parameter vector β, the search of efficient designs based on the Fisher information is sensitive to the assumed values of β used during the design search (Arora and Huber 2001; Ben-Akiva and Lerman 1985; Huber and Zwerina 1996). For the construction of an efficient design that is robust to deviations of the assumed values of β from the true values of β, the use of Bayesian designs that incorporate the uncertainty of the prior knowledge on β has been recommended (e.g., Chaloner and Verdinelli 1995; Sándor and Wedel 2001; Yu, Goos, and Vanderbroek 2009). The distribution $\pi(\beta)$ is set based on prior information. Specifically, for the set of variables X_1 that are believed to be active, we use the prior $\pi(\beta_1) \sim \text{Normal}(\mu, \sigma^2 I)$ for the corresponding vector of parameters β_1, where μ is the mean vector with values away from zero, I is the identity matrix, and σ^2 is the variance. For the set of variables X_2 that are believed to be inactive, we use a prior centered at zero, $\pi(\beta_2) \sim \text{Normal}(0, \sigma^2 I)$, for the parameters in β_2.

Design Characteristics Compared with the Traditional Bayesian D-Criterion

In this section, we examine the use of the proposed Bayesian ϕ-criterion in the construction of efficient designs for a selective choice process and compare the resulting designs (ϕ-designs) with benchmark designs (D-designs) obtained under the Bayesian D-criterion commonly used in the marketing literature (Kessels, Goos, and Vandebroek 2006; Sándor and Wedel 2001).[1] The Bayesian D-criterion focuses on a single objective—efficient estimation of all parameters (β) in the model—and therefore aims to find designs that maximize the following:

$$D = \int \log|I(\beta|X)|\pi(\beta)d\beta. \tag{4}$$

The key difference between the Bayesian D-criterion in Equation 4 and the proposed ϕ-criterion in Equation 3 is that the ϕ-criterion balances between two design objectives. Specifically, given auxiliary self-stated information that distinguishes between active and inactive attributes, the dual objectives of the ϕ-criterion differentially focus on the assumed active attributes while allowing for the possibility that some assumed inactive attributes might indeed be active.

Note that the self-stated prior information distinguishing between active and inactive attributes could also be used to form an informative prior distribution in constructing designs under the Bayesian D-criterion. We compare ϕ-designs obtained using the ϕ-criterion with three versions of D-designs obtained using the Bayesian D-criterion, each with different assumptions for the prior distribution, $\pi(\beta)$. The first is a D-design commonly used in practice where $\pi(\beta)$ is set to be the standard normal distribution for all parameters. This D-design does not make use of the self-stated information, and we call it the D-design with uninformative priors (labeled D_0). The other two D-designs make use of the self-stated information that distinguishes between active and inactive attributes to set the distribution for $\pi(\beta)$. In particular, the D_I (informative mean) D-design uses the active–inactive information in the self-stated data to set the mean of the distribution for $\pi(\beta)$ either at zero for assumed inactive attributes or away from zero for assumed active attributes. The D_S (spike-and-slab) D-design uses the active–inactive information in the self-stated data to set the mean of the distribution for $\pi(\beta)$ away from zero for the assumed active attributes or a mixture distribution for the assumed inactive attributes having an 80% point-mass at zero and 20% drawn from a distribution with a mean away from zero.[2]

Table 1 summarizes the design criteria and core assumptions that characterize each of the designs. Note that all three D-designs use the same single-objective D-criterion, whereas only the ϕ-design uses a compound design criterion. The informative prior D-designs (D_I and D_S) offer alternative approaches to design customization using the same data but without the potential advantages of differential focus across objectives in the ϕ-criterion. The uninformative prior D-design (D_0) is the only design that does not make use of the information in the self-stated data to distinguish between active and inactive elements in the prior; thus, the D_0-design provides a comparison as to the value of including auxiliary data in design construction.

For the benchmarking study, we use a conjoint choice experiment that involves 8 binary attributes and 18 choice sets with two alternatives per choice set. In such a scenario, people may vary in how many of the eight attributes will be active in their choice process. We construct ϕ-designs corresponding to the seven possible variable selection scenarios (one active attribute to seven active attributes) that can be present in the self-stated active–inactive classification data. We use an 80% weight (w = .8) in the compound design criterion on the estimation of the effects of active attributes in the reduced model. In addition, we build a set of informative prior D_I- and D_S-designs that also correspond to the seven possible variable selection scenarios. Each design in the set represents different scenarios for prior information as to which attributes are active or inactive. It is this active–inactive distinction using self-stated prior information from the respondent that gives rise to customized designs. In addition to these three sets of informative prior designs (ϕ, D_I, and D_S), we construct an uninformative prior D_0-design that does not vary in any relationship to the self-stated auxiliary data (i.e., no customization).

We compare all four design types (ϕ, D_0, D_I, and D_S) across a range of contexts and measures. We first examine the efficiency of ϕ-designs relative to the D-designs under specific assumptions. Next, we focus on the characteristics of the different design types to better understand the observed differences in relative efficiency across design types and to illustrate why ϕ-designs offer design performance advantages. Finally, to assess the scope of the observed design advantages, we compare ϕ-designs with all three D-designs under alternative settings and examine design robustness to prior misspecification.

[1]The Bayesian D-criterion has been expressed with or without the logarithmic transformation of the determinant of the Fisher information matrix. We adopt the latter because it is more consistent with the Bayesian theoretic approach (Chaloner and Verdinelli 1995) and brings computational advantages in the design search by making the criterion less sensitive to parameter vectors that lead to very large or very small determinant values (Kessels, Jones, and Goos 2011).

[2]We thank a reviewer for suggesting this alternative approach for using informative priors in D-designs. Although the means and variances of the prior distribution $\pi(\beta)$ could be set by using managers' beliefs or from prior studies (Sándor and Wedel 2001), consistent with previous research on Bayesian designs (e.g., Kessels, Goos, and Vandebroek 2006), in all informative prior designs (ϕ, D_I, and D_S), we assume that $\mu_{active} = 1$ and $\sigma^2_{active} = 1$ in the specification of the prior distribution for the vector of parameters $\pi(\beta_{active})$. For the attributes believed to be inactive, the ϕ-design and the D_I-design both use the same settings for $\pi(\beta_{inactive})$, while the D_S-design uses a mixture for $\pi(\beta_{inactive})$.

Table 1

DESIGN CRITERIA FOR ϕ-DESIGNS AND D-DESIGNS WITH UNINFORMATIVE (D_0) AND INFORMATIVE (D_I, D_S) PRIORS

Design	*Design Criterion*	*Prior for $\pi(\beta_1)$*	*Prior for $\pi(\beta_2)$*	*Objectives in Criterion*	*Active–Inactive Information*
D_0	$\int \log\|I(\beta\|X)\|\,\pi(\beta)d\beta$	N(0, 1)	N(0, 1)	One	No
D_I	$\int \log\|I(\beta\|X)\|\,\pi(\beta)d\beta$	N(1, 1)	N(0, 1)	One	Yes
D_S	$\int \log\|I(\beta\|X)\|\,\pi(\beta)d\beta$	N(1, 1)	.8 × 0 .2 × N(1, 1)	One	Yes
ϕ	$\int\{w[\log\|I(\beta_1\|X_1)\|] + (1 - w)$ $[\log\|I(\beta\|X)\| - \log\|I(\beta_1\|X)\|]\}\,\pi(\beta)d\beta$	N(1, 1)	N(0, 1)	Two	Yes

Notes: In all circumstances, we assume that attribute partworths and the active–inactive attribute classification information are constant across the data-collection event.

Comparison on design efficiency. We compare efficiency of the ϕ-designs relative to D-designs with uninformative (D_0) or informative priors (D_I or D_S) over the distinct variable selection scenarios. Each scenario represents a selective choice process with a different number of inactive attributes. To distinguish from the prior distribution $\pi(\beta)$ assumed in the design construction, we use $\pi_t(\beta)$ to denote the true data-generating process of the parameter vector β, which may deviate from the assumed prior $\pi(\beta)$. To begin, for the parameters corresponding to the inactive attributes, we set $\pi_t(\beta_2) \sim N(0, .5)$ and $\pi_t(\beta_1) \sim N(1, .5)$ for the parameters corresponding to the active attributes. In subsequent analysis, we vary the variances for the data-generating distribution $\pi_t(\beta)$ and examine the relative efficiency of ϕ-designs across a range of settings. We compare the designs using two design efficiency measures:

1. Design efficiency on parameter estimation of all variables (i.e., estimation of β); for a given ϕ-design with model matrix X, the efficiency of the ϕ-design relative to the D_0-, D_I-, or D_S-design with model matrix X^* on estimation of the effects of all variables is defined as

$$(5)\qquad \text{Relative Efficiency on Estimation of } \beta = \frac{\int |I(\beta|X)|^{\frac{1}{m}}\pi_t(\beta)d\beta}{\int |I(\beta|X^*)|^{\frac{1}{m}}\pi_t(\beta)d\beta}.$$

2. Design efficiency on parameter estimation of the active variables (i.e., estimation of β_1); specifically, for a given ϕ-design with model matrix X, the efficiency of the design relative to the D_0-, D_I-, or D_S-design with model matrix X^* on estimation of the effects of active variables is defined as

$$(6)\qquad \text{Relative Efficiency on Estimation of } \beta_1 = \frac{\int |I(\beta_1|X_1)|^{\frac{1}{m_1}}\pi_t(\beta_1)d\beta_1}{\int \left|I\left(\beta_1|X_1^*\right)\right|^{\frac{1}{m_1}}\pi_t(\beta_1)d\beta_1}.$$

We evaluate these measures over 1,000 realizations of the distribution $\pi_t(\beta)$. For both measures, a value greater than 1 suggests that the ϕ-design with model matrix X is more efficient than the D_0-, D_I-, or D_S-designs under comparison, while a value less than 1 suggests the opposite.

The results presented in Table 2 are the relative efficiency measures computed when the set of inactive attributes in the true data-generating process ("variable selection pattern") is concordant with that assumed by the active–inactive prior used in the construction of the designs. For example, row 5 of Table 2 shows the scenario that assumes a variable selection pattern with three inactive and five active attributes. When evaluated under the data-generating process with a concordant variable selection pattern, the ϕ-design is 40% more efficient than the D_0-design, 6.6% more efficient than the D_I-design, and 5.2% more efficient than the D_S-design on estimation of the effects of all variables. On average, across all the scenarios, ϕ-designs are 30.2% more efficient than the uninformative prior D_0-design and 3.7% and 6.1% more efficient than the corresponding D_I- and D_S-designs, respectively, on the estimation of the effects of all variables. The efficiency advantage of ϕ-designs relative to D-designs is even more striking considering the relative efficiency on estimation of the effects of the active variables shown in the rightmost three columns of Table 2.

Note that the efficiency gains for the ϕ-designs relative to the D_0-design come from two sources: informative prior means and differential emphasis across attributes facilitated by the compound design criterion. In contrast, the efficiency gains for the ϕ-designs relative to the D_I- and D_S-designs come from the latter source only. This is because the D_I-and D_S-designs also utilize the same prior information but lack the flexibility for

Table 2

DESIGN EFFICIENCY OF ϕ-DESIGNS RELATIVE TO D-DESIGNS WITH UNINFORMATIVE (D_0) AND INFORMATIVE (D_I, D_S) PRIORS

	Relative Efficiency on Estimation of All Variables (β)			*Relative Efficiency on Estimation of Active Variables (β_1)*		
Design Scenario	*ϕ/D_0*	*ϕ/D_I*	*ϕ/D_S*	*ϕ/D_0*	*ϕ/D_I*	*ϕ/D_S*
1 Active: 7 Inactive	1.047	.998	1.003	1.286	1.800	3.600
2 Active: 6 Inactive	1.104	1.036	1.072	1.109	1.118	1.974
3 Active: 5 Inactive	1.214	1.026	1.053	1.081	1.010	1.455
4 Active: 4 Inactive	1.422	1.104	1.161	1.313	1.143	1.544
5 Active: 3 Inactive	1.400	1.066	1.052	1.347	1.066	1.143
6 Active: 2 Inactive	1.496	1.009	1.051	1.413	1.018	1.101
7 Active: 1 Inactive	1.431	1.020	1.032	1.375	1.021	1.046
Average	1.302	1.037	1.061	1.275	1.168	1.695

Notes: Relative efficiency in data context with variable selection concordant to design where $\pi_t(\beta_{active}) \sim N(1, .5)$ and $\pi_t(\beta_{inactive}) \sim N(0, .5)$.

differential focus across active and inactive attributes through the two objectives of the compound design criterion. To better understand why ϕ-designs offer advantages relative to comparable D-designs, next we examine the characteristics of the designs and illustrate the design differences using a stylized example.

Design characteristics. We follow extant research and focus on four design characteristics: level overlap, level balance, orthogonality, and utility balance. Level overlap is repeated occurrences of an attribute level within a choice set, and we measure it by calculating the average overlap rate (Sándor and Wedel 2002). Level balance refers to equal occurrences of each level within an attribute. We measure the deviation from level balance by the mean absolute deviation (MAD) of level frequency (Liu and Arora 2011). A design with higher MAD of level frequency has lower level balance. Orthogonality refers to equal occurrences of pairs of attribute levels. An orthogonal design has zero correlation; thus, we measure deviation from orthogonality using the average absolute pairwise correlation of the mean-adjusted columns in model matrix X (Liu and Arora 2011). Utility balance refers to equally attractive alternatives within choice sets, and we measure it with cumulative entropy as

$$-\sum_{s=1}^{S}\int_{R}\left[\sum_{j=1}^{J}\left\{p_{sj}(X_s,\beta)\ln\left[p_{sj}(X_s,\beta)\right]\right\}\right]\pi_t(\beta)d\beta,$$

calculated using R = 1,000 samples from $\pi_t(\beta)$. A design with higher entropy is more utility balanced, but designs with the highest entropy are not necessarily the most efficient (Kessels, Goos, and Vandebroek 2006).

Table 3 presents a summary of the characteristics of ϕ-designs and D-designs with uninformative (D_0) and informative (D_I, D_S) priors, averaged across all seven variable selection scenarios. Comparing ϕ-designs with the D-design with an uninformative prior (D_0), we find that the ϕ-designs have higher level overlap for active attributes but lower overlap for inactive attributes, lower level balance for active attributes but higher level balance for inactive attributes, lower orthogonality, and higher utility balance. Consistent with findings from prior research, higher level overlap is associated with lower level balance (Liu and Arora 2011). Because of the relative weight placed on the active attributes in the design criterion, high overlap of active attributes is associated with low orthogonality and high utility balance of the entire design. These results make intuitive sense because of the repetitions of the same attribute level within a choice set when there is level overlap.

The differences on level overlap between the ϕ- and D_0-designs can be explained by the assumed variation in attribute levels used during design construction. Efficient designs tend to maximize variation in attribute levels to maximize information on the estimators of parameters β (Sándor and Wedel 2002). If no variation is assumed among attribute levels during design construction (i.e., the prior $\pi(\beta)$ is assumed to be a point mass at 0), then the only source of variation comes from varying the attribute levels in choice sets—leading to designs with minimal level overlap (Huber and Zwerina 1996). Compared with the D_0-design with an uninformative prior mean of zero (i.e., no expected differences between attribute levels), ϕ-designs assume more variation in the attribute levels for the active attributes through the informative prior with a mean

Table 3

DESIGN CHARACTERISTICS OF ϕ-DESIGNS AND D-DESIGNS WITH UNINFORMATIVE (D_0) AND INFORMATIVE (D_I, D_S) PRIORS

Design Characteristic	ϕ	D_0	D_I	D_S
Level overlap for active variables	.246	.202	.310	.437
Level overlap for inactive variables	.071	.214	.177	.008
MAD of level frequency for active variables	1.429	1.286	1.500	2.714
MAD of level frequency for inactive variables	1.000	2.214	1.536	.143
Deviation from orthogonality (Average absolute pairwise correlation in model matrix X)	.081	.072	.083	.080
Utility balance (Cumulative entropy for model matrix X and $\pi_t(\beta)$)	3.978	3.794	4.275	4.730

Notes: averages across all design sparsity scenarios where for utility balance the variable selection pattern is concordant to the design and $\pi_t(\beta_{active}) \sim N(1, .5)$ and $\pi_t(\beta_{inactive}) \sim N(0, .5)$.

away from zero (i.e., expected differences between levels of active attributes). Therefore, ϕ-designs do not require as much variation of attribute levels in choice sets and thus lead to higher overlap rates for the active attributes. In contrast, although both the D_0- and ϕ-designs use the same prior for the inactive attributes, it only takes effect in the second component of the ϕ-criterion. In the first component of the ϕ-criterion, in which the emphasis is on the estimation of the effects of only the active attributes in the reduced model, the prior on the inactive attributes does not have an effect. Therefore, the "effective" variation assumed for inactive attributes in the construction of ϕ-designs is much smaller than in the D_0-design. Thus, ϕ-designs require more variation of attribute levels in choice sets—leading to lower level overlap for the inactive attributes.

Next, we compare ϕ-designs with D-designs with informative priors (D_I and D_S). Because level overlap is associated with the other three characteristics consistently, we focus the discussion on level overlap only. For active attributes, ϕ-designs have lower level overlap than both D_I- and D_S-designs. For inactive attributes, ϕ-designs have lower level overlap than D_I-designs but higher level overlap than D_S-designs. While the lower overlap for the inactive attributes compared with D_I-designs can be explained by "effective" assumed variation for inactive attributes, the other observed differences in level overlap are due to the differential emphasis that the ϕ-criterion places on the two design objectives. Specifically, the greater weight on efficient estimation of active attributes leads to lower overlap for active attributes in ϕ-designs compared with the D_I- and D_S-designs.

To better understand this, consider a stylized example with two binary attributes A and B, in which the prior information indicates that A is likely to be active and B is likely to be inactive. A choice set with two alternatives such as (A_1B_1, A_1B_2) or (A_2B_1, A_2B_2), in which there is overlap on the levels of attribute A, makes it easier to detect any difference between the two levels of the inactive attribute B than choice sets with no level overlaps in the active attribute such as (A_1B_1, A_2B_2) or (A_2B_1, A_1B_2). Thus, level overlap in the active attribute A helps with the detection of the effects of attribute B, which is stated to be inactive but may turn out to be active. However, choice sets with level overlap in A do not reveal any information about the difference between the two levels of A. There must be some choice sets with no level overlap in the

Table 4
DESIGN EFFICIENCY OF ϕ-DESIGNS RELATIVE TO D-DESIGNS WITH UNINFORMATIVE (D_0) AND INFORMATIVE (D_I, D_S) PRIORS IN DIFFERENT SETTINGS FOR $\pi_t(\beta)$

	Relative Efficiency on Estimation of All Variables (β)			*Relative Efficiency on Estimation of Active Variables* (β_1)		
Setting for $\pi_t(\beta)$	ϕ/D_0	ϕ/D_I	ϕ/D_S	ϕ/D_0	ϕ/D_I	ϕ/D_S
$\pi_t(\beta) \sim N(0,1)$	.962	1.022	1.105	.955	1.149	1.691
$\pi_t(\beta_{active}) \sim N(1,1)$ $\pi_t(\beta_{inactive}) \sim N(0,1)$	1.179	1.020	1.065	1.194	1.162	1.686
$\pi_t(\beta_{active}) \sim N(1,1)$ $\pi_t(\beta_{inactive}) \sim .8(0) + .2[N(1,1)]$	1.275	1.019	.879	1.194	1.162	1.683
$\pi_t(\beta_{active}) \sim N(1,.5)$ $\pi_t(\beta_{inactive}) \sim N(0,.5)$	1.302	1.037	1.061	1.275	1.168	1.695

Notes: Average relative efficiency across all design sparsity scenarios in different data settings for $\pi_t(\beta)$ and where the variable selection pattern is concordant to the design.

active attribute A to efficiently estimate that effect. Therefore, because the ϕ-criterion places more weight on the efficient estimation of the effects of the active attributes, it makes intuitive sense that the ϕ-designs have lower level overlap for the active attributes than the D-designs with informative priors (D_I and D_S). Conversely, the lesser weight that the ϕ-criterion places on the detection of the effects of the inactive attributes explains why the ϕ-designs have higher level overlap for the inactive attributes than the D_S-designs.

In summary, the design characteristics comparisons indicate that the benefits of ϕ-designs originate not only from the informative priors $\pi(\beta)$ on the active–inactive status of attributes but also from the differential emphasis on the two objectives in the ϕ-criterion. This differential emphasis leads to choice sets with alternatives that focus more on the comparisons of active attributes, making ϕ-designs more efficient on the estimation of the active attributes while retaining the ability to detect the effects of the attributes stated as inactive that may turn out to be active. Cumulatively, the results provide evidence in favor of using the ϕ-criterion in constructing efficient designs, with the caveat that the reported efficiency advantages assume that the variable selection pattern in the actual data is concordant with that assumed in the design construction. What happens when the actual variable selection pattern deviates from the prior assumptions? Next, we consider three interrelated elements of that question: (1) impact of alternative variance settings for the data distribution $\pi_t(\beta)$, (2) design robustness to sparsity misspecification, and (3) specification of the weight w in the compound design criterion.

Variance settings for the data distribution $\pi_t(\beta)$. To establish the potential bounds of ϕ-design efficiency advantages, we evaluate all designs under the "best-case scenario" for each particular design, which is the data-generating process consistent with the assumed prior (i.e., $\pi(\beta) = \pi_t(\beta)$). The results in Table 4 show that when the data-generating process is consistent with the spike-and-slab mixture prior, D_S-designs are relatively more efficient than ϕ-designs on estimation of all variables. Likewise, when the data-generating process is N(0,1), the D_0-design with uninformative priors is relatively more efficient than ϕ-designs. When the data-generating process is consistent with the prior used in building both the ϕ- and D_I-designs, such that $\pi_t(\beta_{active}) \sim N(1, 1)$ and $\pi_t(\beta_{inactive}) \sim N(0, 1)$, the ϕ-designs are relatively more efficient.

We examine the efficiency of ϕ-designs relative to each of the D-designs across a grid of variance settings ranging from .1 to 1 for both the active and inactive attributes in $\pi_t(\beta)$. As such, for each D-design, we obtain a 10 × 10 matrix of relative efficiency measures on the estimation of the effects of all variables. On average, across all the variance settings, ϕ-designs are 30.5% more efficient than D_0-designs, 3.1% more efficient than D_I-designs, and 5.3% more efficient than D_S-designs. The surfaces plotted in Figure 1 show the efficiency of ϕ-designs relative to the three D-designs across all 100 variance settings. We find that ϕ-designs are consistently more efficient than both D_0 and D_I designs in 100% of the variance settings and more efficient than D_S-designs in 84% of the variance settings. These results provide strong evidence for the efficiency advantages of ϕ-designs over D-designs, under the assumption that the variable selection pattern in the data is concordant to that in the self-explicated prior. Next, we examine whether and how the results may change when the prior is misspecified.

Design robustness to sparsity misspecification. When the prior information about active and inactive variables used in building a design deviates from the actual variable selection pattern in the data, sparsity is misspecified. In our study, with eight binary attributes, there is just one of 2^8 possible realizations of the active–inactive variable selection pattern in which the data matches the prior assumed for design construction. Up to this point, design performance under such concordance has been the context for analysis, but we next consider design robustness to sparsity misspecification for the remaining $2^8 - 1$ cases that vary in the extent of data–prior discordance.

For each of the seven variable selection scenarios used in design construction, we compute the efficiency of ϕ-designs relative to all three D-designs (D_0, D_I, and D_S) across all of the of 2^8 possible outcomes of active–inactive data over 1,000 realizations of the four settings of $\pi_t(\beta)$ shown in Table 4. For each variable selection scenario, the 2^8 possible active–inactive outcomes can be organized according to the extent of discord between the data and the prior—ranging from none wrong to all wrong. For example, there is just one case in which the data and the prior are in complete concordance ("zero wrong"), and there are

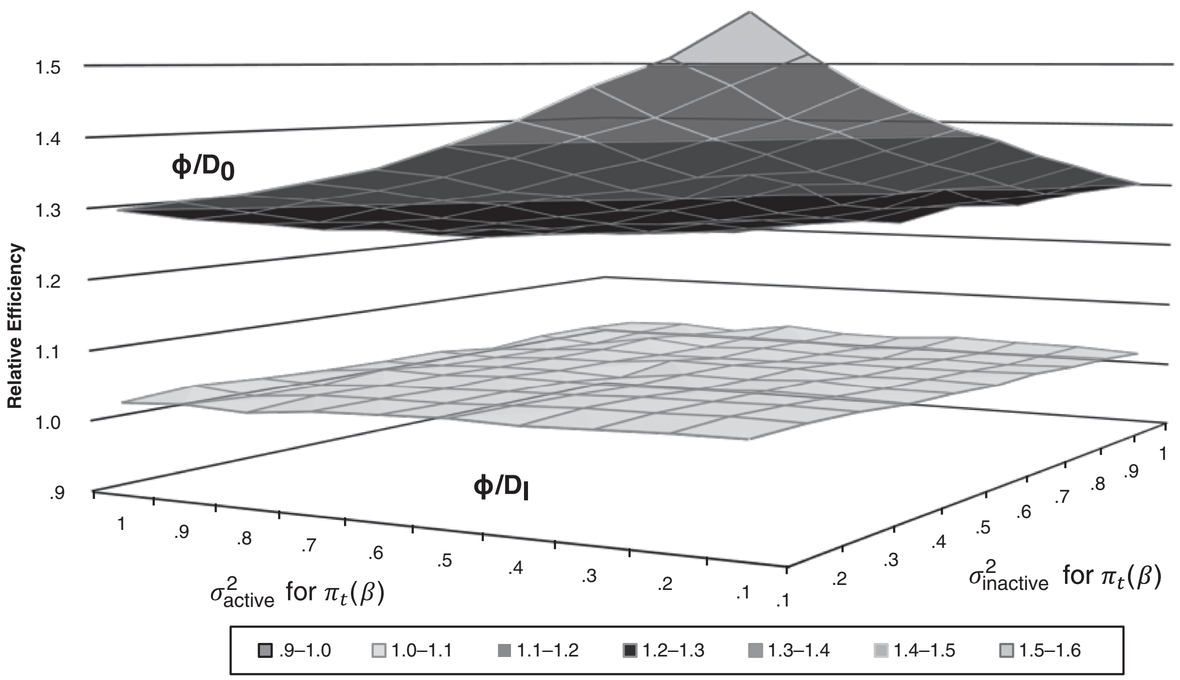
$$\binom{8}{2} = 28$$

cases in which there is discord between the data and the prior on the active–inactive status of two attributes ("two wrong").

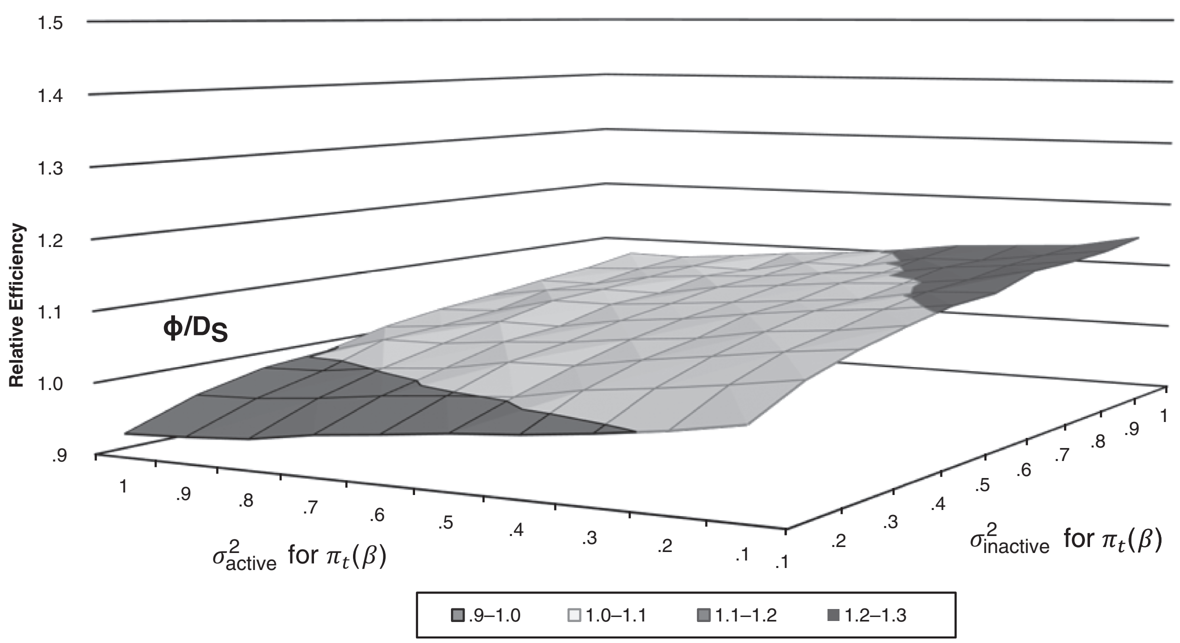
Figure 1

DESIGN EFFICIENCY IN ESTIMATION OF ALL VARIABLES OF ϕ-DESIGNS RELATIVE TO D-DESIGNS WITH UNINFORMATIVE (D_0) AND INFORMATIVE (D_I, D_S) PRIORS IN DIFFERENT VARIANCE SETTINGS (σ^2_{active} AND $\sigma^2_{inactive}$) FOR $\pi_t(\beta)$

A: Design Efficiency of ϕ-Designs Relative to D-Designs with Uninformative (D_0) and Informative (D_I) Priors



B: Design Efficiency of ϕ-Designs Relative to D-Designs with Informative (D_S) Priors



Notes: Figure surfaces denote average relative efficiency across all design sparsity scenarios at positions of the grid of variance settings for $\pi_t(\beta)$ where the variable selection pattern is concordant to the design (the Web Appendix reports tabular values).

We organize the relative efficiency results according to the extent of discord. The full results for all four base settings of $\pi_t(\beta)$ are reported in the Web Appendix, while a graphical representation of the relative efficiency under the spike-and-slab data setting for $\pi_t(\beta)$ appears in Figure 2. These results enable us to make prescriptive statements as to where ϕ-designs offer benefits relative to D-designs.

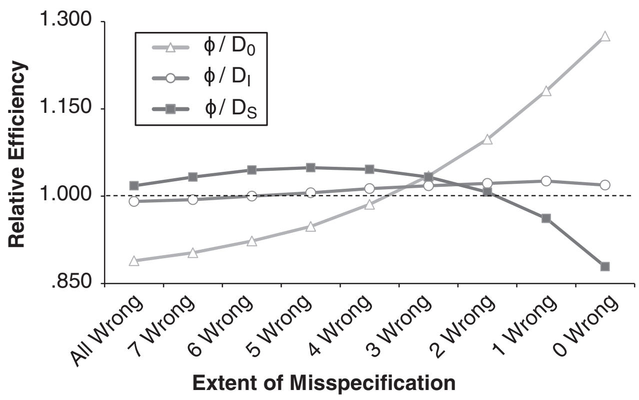
As Figure 2 shows, ϕ-designs become relatively less efficient than D_0-designs when at least half of the prior is mismatched to the data ($\geq$4 wrong). Thus, if the relationship between the active–inactive self-stated prior and the variable selection pattern in the data is no better than chance, then D_0-designs with uninformative priors are most efficient. This result holds across all four $\pi_t(\beta)$ settings, demonstrating that if the active–inactive prior is not informative, then no design making use of that prior will be more efficient than a D_0-design. However, Figure 2 also shows that if the relationship between the active–inactive self-stated prior and the variable selection pattern in the data is perfect, or nearly so, and the distribution of $\pi_t(\beta)$ is a mixture distribution, then spike-and-slab prior D_S-designs are reasonable options. However, the tightness of the mixture distribution for the inactive attributes makes the D_S-design less robust to sparsity misspecification. The D_S-design efficiency advantage disappears as the discrepancy between the data and the prior increases, or as the distribution of $\pi_t(\beta)$ deviates from a spike-and-slab mixture distribution. Overall, if the fidelity between the prior specification and the data reality lies on the range between random chance and perfect correspondence, the results show that ϕ-designs are the approach best suited for the context.

Impact of criterion weight (w) *on design efficiency.* In Equation 3, the weight (w) must be set a priori to balance the objectives in the compound design criterion. Higher weight values place more focus on the efficient estimation of the effects of active variables in the reduced model but also lessen the design's ability to accommodate sparsity misspecification of the prior. At the extreme, a weight of 1 results in a design criterion focused entirely on the reduced model, without any possibility for inactive variables that end up being active. Table 5 shows the efficiency of ϕ-designs built using weights of .6, .7, .8, and .9 relative to each of the three D-designs (D_0, D_I, and D_S). Importantly, the results show that although higher-weight designs are indeed relatively more efficient in estimating the effects of active attributes in the reduced model, the superiority of ϕ-designs relative to the different D-designs holds true across the range of weight settings. In our empirical study, we use a weight w = .8. If prior data (e.g., through pilot studies) are available on the degree of concordance between the self-stated active–inactive status of attributes and that reflected in actual choices, such information can be used to set the weight w to achieve higher efficiency of the resulting designs.

In summary, our comparative study shows that using prior information to create ϕ-designs customized to a specific pattern of variable selection leads to efficiency gains relative to D-designs with uninformative (D_0) or informative (D_I and D_S) priors. We demonstrate this efficiency advantage across a range of weight settings for the compound design criterion and under different variance settings of the data distribution. The results give clear direction for design construction under different data contexts. Specifically, an

Figure 2

DESIGN EFFICIENCY OF ϕ-DESIGNS RELATIVE TO D-DESIGNS WITH UNINFORMATIVE (D_0) AND INFORMATIVE (D_I, D_S) PRIORS UNDER DIFFERENT DEGREES OF MISSPECIFICATION WHERE THE SETTINGS FOR $\pi_t(\beta)$ ARE CONSISTENT WITH THE SPIKE-AND-SLAB MIXTURE PRIOR



Notes: Each line represents the relative efficiency of ϕ-designs compared with different D-designs over all 2^8 possible realizations of the data, averaged across all seven sparsity designs. The number of realizations of misspecification that correspond to each position on the x-axis varies because, for example, there is just one circumstance in which the prior and the data can be in 100% concordance, but there are (28) possible ways to have the prior misspecified in two places.

uninformative D_0-design does not require self-stated data and is easy to implement without the need for customization. If the active–inactive self-stated prior has no relationship to the actual variable selection pattern in the data, then the analyst is better off using an uninformative D_0-design—forgoing the added data collection and design customization effort. The D-criterion designs with informative priors (D_I and D_S) have data requirements similar to the ϕ-design. If auxiliary data are to be collected and self-stated prior information is to be used in the construction of designs customized to a particular variable selection pattern, then the results of our comparative study across an array of data circumstances show that ϕ-designs dominate informative prior D-designs in all but a few specific circumstances. Hereinafter, we focus on a comparison between the ϕ-design (i.e., customization) and the traditional D_0 design (i.e., no customization), which serves as the reference benchmark.

Previously, we evaluated designs using measures that rely on the design matrix and the assumed data distribution. As we implement a data collection and customization strategy, designs can also be evaluated by simulating or collecting choice data, fitting a model to that data, and then comparing the model parameter estimates. For example, performance measures focused on accuracy (e.g., expected mean square errors of the parameters), predictive validity (e.g., out-of-sample hit rate), or managerial inference (e.g., willingness to pay) can all be evaluated this way. In the next section, we introduce the heterogeneous variable selection model that we use to assess the performance of ϕ-designs both in simulations and in our empirical study.

EFFICIENT ANALYSIS OF DATA ARISING FROM A SELECTIVE CHOICE PROCESS

Having shown the design benefits of using the ϕ-criterion, we shift focus to analysis and present our approach to modeling a selective choice process that is characterized by possible inactive attributes, self-stated priors, and customized designs. Just as the ϕ-criterion leverages auxiliary information in creating a design customized to a specific sparsity scenario, in this section we extend variable selection models to accommodate this same self-stated data. First, we use a simulation to demonstrate the benefits of analysis for a selective choice process that incorporates auxiliary data both when building customized designs and when estimating heterogeneous variable selection models. In the subsequent empirical application, we collect self-stated active–inactive priors, build customized designs, observe choices, and show the substantive implications of efficiently analyzing data arising from a selective choice process. We begin by specifying the basic utility-maximizing model of individual choice.

For individual i (i = 1, ..., N) evaluating j (j = 1, ..., J) alternatives within choice set s (s = 1, ..., S), the task is to choose the alternative (j′) with the highest utility using the choice rule:

$$y_{is} = j' \text{ if } V_{isj'} + \varepsilon_{isj'} \geq V_{isj} + \varepsilon_{isj} \quad \forall j. \tag{7}$$

The deterministic component of utility ($V_{isj} = x'_{isj}\beta_i$) of each alternative j for individual i making choice s has a design vector (x_{isj}) indicating the combination of attribute levels that describes the choice alternative and individual sensitivity to those variables (β_i). When combined with the stochastic component of utility, $\varepsilon_{isj} \sim EV(0, 1)$, the choice probability takes the form $\Pr(y_{is} = j') = \exp(V_{isj'})/\sum_{j=1}^{J}\exp(V_{isj})$.

Heterogeneous Variable Selection Model

The basic premise of a selective choice process is that certain attributes do not matter to a decision maker. These inactive attributes do not contribute to an alternative's expected utility; thus, one goal of an individual-level model for a selective choice process is to distinguish between active and inactive attributes. We capture variable selection at the individual level for M predictors in a model using augmented data parameters τ_{im} for all m predictors, where $\tau = \{c, 1\}$ (Gilbride, Allenby, and Brazell 2006). The value c is set close to zero (c = .01) to indicate that the variable is inactive when computing expected utility, while a value of $\tau = 1$ indicates that the attribute associated with the variable is active in decision making. The attribute associated with parameter β_{im} is active ($\tau_{im} = 1$) with probability θ_m. In this variable selection model, the choice rule remains as expressed in Equation 7, but the deterministic portion of utility, $V_{isj} = x'_{isj}(\tau'_i\beta_i)$, has a different interpretation because the variable selection parameter (τ_{im}) can effectively set some of the attribute sensitivity parameters (β_{im}) in the utility function to zero.

The variable selection parameters (τ_{im}) are indicators of the active–inactive status of a variable in a person's deterministic utility function. These parameters $\{\tau_i\}$ form the diagonal elements of an M × M diagonal matrix $C_{\tau i}$ that is used to modify the random-effects distribution of heterogeneity as $\beta_i^* = C_{\tau i}\beta_i \sim N(C_{\tau i}\overline{\beta}, C_{\tau i}\Sigma C_{\tau i})$. The general form of the heterogeneity distribution and its relationship to the active–inactive indicators can be expressed as $\beta_i = C_{\tau i}^{-1}\beta_i^* \sim N(\overline{\beta}, \Sigma)$. In this formulation, the baseline random-effects logit model is a special case of the model where all variable selection indicators τ_{im} are fixed at 1 and $C_{\tau i}$ is an identity matrix.

Table 5

DESIGN EFFICIENCY OF ϕ-DESIGNS WITH DIFFERENT WEIGHT SETTINGS (w) RELATIVE TO D-DESIGNS WITH UNINFORMATIVE (D_0) AND INFORMATIVE (D_I, D_S) PRIORS

	Relative Efficiency on Estimation of All Variables (β)			*Relative Efficiency on Estimation of Active Variables (β_1)*		
Setting for (w)	ϕ/D_0	ϕ/D_I	ϕ/D_S	ϕ/D_0	ϕ/D_I	ϕ/D_S
w = .6	1.308	1.034	1.058	1.186	1.056	1.494
w = .7	1.307	1.036	1.061	1.241	1.123	1.681
w = .8	1.302	1.037	1.061	1.275	1.168	1.695
w = .9	1.321	1.045	1.070	1.307	1.191	1.723

Notes: Average relative efficiency across all design sparsity scenarios in different data settings for (w), where the variable selection pattern is concordant to the design and where $\pi_t(\beta_{active}) \sim N(1, .5)$ and $\pi_t(\beta_{inactive}) \sim N(0, .5)$.

Incorporating Auxiliary Information

When collecting data, respondents can provide information about the status of variables in their decision making independent of the choice task. This information can help identify which attributes do or do not matter to the decision maker and can be used as prior information both in design customization and in estimating the latent variable selection parameters. Previous research has shown the value of incorporating auxiliary information into models in other contexts (e.g., Aribarg, Arora, and Kang 2010; Arora, Henderson, and Liu 2011; Lenk and Orme 2009), and its relevance in modeling a selective choice process is important for two reasons. First, although individual respondents may not use the terminology of "active–inactive attributes" and "zero utility," it is quite easy to collect data in which, given a range domain, respondents organize attributes into two groups: those that "do matter" and those that "do not matter." Rather than relying solely on data collected in the choice task itself, the aforementioned binary attribute classification can be used to help uncover variables' latent active–inactive status. Second, by definition, inactive variables have little impact on utility; thus, detecting their effects on the likelihood requires nuance. As such, leveraging auxiliary data is potentially useful for variable selection models and is as yet an unaddressed topic.

Given the auxiliary binary active–inactive attribute classification data (effects coded and denoted as z_{im}), the probability that variable m is active can depend on that binary grouping such that $\Pr(\tau_{im} = 1) = \theta_{im}$ where $\theta_{im} = \exp(\gamma_m z_{im})/[1 + \exp(\gamma_m z_{im})]$. We estimate the (M × 1) vector γ representing the relationship between the self-stated prior data and the model-based variable selection behavior instead of the (M × 1) vector θ, which does not use the auxiliary information. An advantage of this approach to incorporating auxiliary data into the model is the ability to assess the fidelity of the self-stated active–inactive data to variable selection behavior inferred from the model. If the value of γ_m for a given attribute is zero, then the information contained in z_m for that attribute has little bearing on the pattern of variable selection manifest in the choice data. Instead, if the value of γ_m is greater than zero,

Table 6
DESIGN PERFORMANCE METRICS FROM SIMULATION STUDY BY DESIGN TYPE

	D-Design	*ϕ-Design*
EMSE(β|X)	.226	.153[a]
$\text{EMSE}_{\text{prob}}(\beta\|X)$	.078	.064[a]
Out-of-sample hit rate	.635	.657[a]

[a]Customized ϕ-designs are superior to D-designs in >95% of realizations.

the self-stated attribution of active–inactive status is relevant to the variable selection behavior inferred from the choice data.

In our approach to design and analysis for a selective choice process, the auxiliary self-stated data serves dual purposes: first, as prior information on the variable selection pattern in the customized ϕ-design construction, and second, as prior information (z_{im}) used in estimating the heterogeneous variable selection model. Without the former, there would not be designs customized to individual-specific patterns of self-stated variable selection, and without the latter, the linkage between the auxiliary data and the modeled behavior would be incomplete. Next, we describe the approach to design customization that we utilize in our empirical applications.

Design Customization

The ϕ-criterion requires that the active variables (X_1) be identified as distinct from the inactive variables (X_2). The self-stated auxiliary data distinguishing between attributes that "do matter" and "do not matter" serve this purpose when building individual-specific designs customized to a particular variable selection scenario. In the current design context, in which eight attributes each have two levels, there are seven distinct patterns of variable selection—ranging from one inactive attribute to seven inactive attributes. These seven patterns represent the source designs that have been the focus of our analysis thus far. Altogether, assuming that there is at least one active attribute, there are $2^8 - 1$ possible combinations of active–inactive self-stated attribute status values. For the extreme case in which all attributes are active, there is no variable selection. Because all attributes in the design have the same number of levels, the elements in X are column-wise exchangeable. As such, in implementing the customization strategy, we reorder the columns in X such that the location of the columns in X_1 matches the pattern of sparsity in the auxiliary data. For example, if there are five active and three inactive attributes, there are

$$\binom{8}{5} = 56$$

possible locations for the active attributes, the designs for which can all be built using the five active and three inactive attribute source ϕ-design simply by reordering the column elements of X.[3]

This approach to design customization suggests that the following steps be used when implementing an empirical study.

Table 7
ATTRIBUTES AND LEVELS USED IN EMPIRICAL CHOICE STUDY

Attribute	*Levels*
Screen size	7 inches or 10.1 inches
Storage	16 GB or 32 GB
Camera resolution	5.0 megapixel or 8.0 megapixel
Price	\$349 or \$449
Processor type	Dual core or quad core
Memory technology type	DDR2 or LPDDR
Display technology type	TFT LCD or IPS LCD
Warranty type	Standard or extended

First, the source designs representing the potential patterns of variable selection can be prebuilt using the ϕ-criterion and stored in a design library. Second, in the initial stage of the choice study, the "does matter" and "does not matter" auxiliary self-stated attribute classification data (z) can be collected from respondents. Third, the variable selection source design in the library with the same number of active attributes as expressed in the self-stated data can be customized by exchanging the columns of X such that the location of X_1 matches the self-stated data for the active attributes. In our empirical application, customization is done instantaneously using JavaScript we wrote and embedded in our survey instrument.[4]

Assessing Accuracy and Predictive Validity of Customized Designs

In this section, using simulation, we examine the relative performance of designs built using the ϕ-criterion and customized to an individual variable selection pattern using the aforementioned strategy. We contrast these customized designs to the benchmark D-design with uninformative priors (D_0) commonly used in practice. In the simulation, we assume there is one group of individuals (N = 400) with heterogeneous preferences and the manifest pattern of variable selection. These pseudo-individuals make choices in both a customized ϕ-design and a no-customization D_0-design context. We compare design accuracy and predictive validity between the two design conditions using the heterogeneous variable selection model.

For the simulated data, we assume a heterogeneous distribution of variable selection behavior, with 40% of pseudo-individuals having one inactive attribute, 30% having two inactive attributes, 20% having three inactive attributes, and 10% having four inactive attributes. We place no restrictions as to which attributes can be active or inactive and randomly assign the variable selection vector τ_i from a location set defined over the number of active attributes. As such, the extent of sparsity varies across individuals and the pattern of sparsity varies across attributes. The preference parameters (β_i) are generated according to $(\beta_{\text{inactive}}|\tau_{im} = c) \sim N(0, .1)$ for the inactive attributes and $(\beta_{\text{active}}|\tau_{im} = 1) \sim N(1, .1)$ for the active attributes. The variable selection indicators (τ_i) are connected to the self-stated "does matter" and "does not matter" indicators (z_i) through $\Pr(z_{im} = 1) = \exp(\gamma\tau_{im})/[1 + \exp(\gamma\tau_{im})]$ where $\gamma = 1.4$ such that there is a 20% possibility that attributes self-stated as inactive in the prior turn out to be truly active in the data.[5]

[3]This approach to customization can also scale to larger design problems. For example, if there were 12 binary attributes, the design library would require the 11 variable-selection pattern source designs plus a reference all-active design rather than 4,095 (= $2^{12} - 1$) designs. When the number of levels is unbalanced across attributes, the source design library would need to be larger and the column swapping would need to be conditional on the level count, but the computational advantage of avoiding a full design space search is notable.

[4]The Web Appendix provides an illustration of the Java coding used in our web-based survey instrument.

[5]We investigated other settings for the inactive proportions as well as the probability of concordance (see the Web Appendix) and find similar results.

Figure 3

SELF-STATED ATTRIBUTE CLASSIFICATION TASK

A **tablet computer** is a one-piece mobile computer with a touchscreen. In the next series of questions you will be shown different tablet computer configurations and will be asked to choose the option that is most attractive to you.

Often, when making choices, there may be some attributes that really matter to you, while the value of some other attributes really don't matter to you. Based on your own preferences, please classify the following tablet computer attributes as: "this attribute **does matter** to me in my decision making" or "this attribute **does not matter** to me in my decision making".

	This attribute **DOES MATTER** to me	This attribute **DOES NOT MATTER** to me
Screen Size	○	○
Storage Capacity	○	○
Camera Resolution	○	○
Price	○	○
Processor Type	○	○
Memory Technology Type	○	○
Display Technology Type	○	○
Warranty Type	○	○

All parameters and variable selection indicators for each pseudo-individual (β_i, τ_i, z_i) as well as the random errors from the extreme value distribution remain the same across the two design conditions—the only thing that differs across conditions is the choice design. In the no-customization condition, all pseudo-individuals are assigned design matrix X^*, from the benchmark D_0-design with uninformative priors, regardless of that pseudo-individual's corresponding self-stated auxiliary data (z_i). In the customization condition, the auxiliary information in z_i is used to create a customized design matrix X_i based on the appropriate source ϕ-design. Given the design matrix, choices are simulated according to the decision rule expressed in Equation 7.

We estimate the aforementioned Bayesian heterogeneous variable selection model for the two design condition data sets and evaluate the two design strategies using parameter recovery and predictive validity metrics. The information in the self-stated auxiliary data z_i is used in model estimation for both the no-customization and customization conditions; thus, any differences between the two conditions are a result of the different design strategies and not the result of different models estimated. We generate 100 replicates of the simulation scenario (i.e., the model was estimated for each design condition for 100 different realizations of the parameters) and conduct inference over these replications.

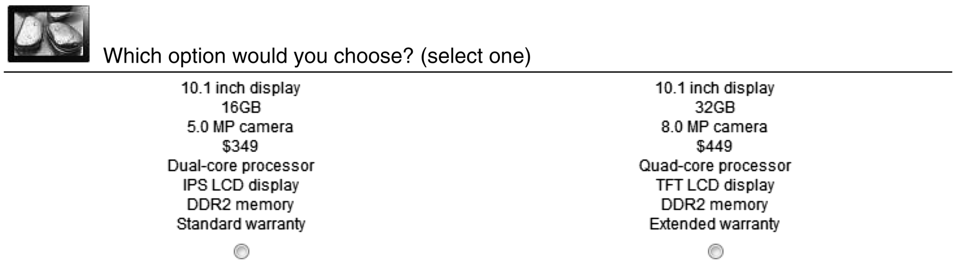
We compare the design strategies based on the expected mean squared errors (EMSEs) in parameter estimates and predicted probabilities (Kessels, Goos, and Vandebroek 2006)[6] as well as the out-of-sample hit rate. For the out-of-sample hit rate, each pseudo-individual in the simulation made choices in two simple holdout tasks generated using SAS OPTEX (Kuhfeld, Tobias, and Garratt 1994; see the Web Appendix). There was no customization of these holdout tasks across pseudo-individuals or across design conditions. Because the holdout tasks are the same across the two design conditions, any performance difference is due to more accurate recovery of the underlying parameters. The lower EMSE values and higher out-of-sample hit rate in Table 6 show that the customized ϕ-designs are consistently better than the benchmark D-design at more faithfully recovering true underlying preferences and achieving better predictive validity. Next, we implement the design customization strategy in a choice experiment study for tablet computers and demonstrate the benefits of ϕ-designs in an applied context.

CUSTOMIZED CONJOINT CHOICE EXPERIMENT

The primary purpose of the conjoint choice experiment is to demonstrate the value of a customized ϕ-design strategy in an

[6]The Web Appendix provides details for the calculation of the EMSEs.

Figure 4
CHOICE TASK FROM EMPIRICAL STUDY



empirical setting. We chose a tablet computer as the focal product because it is a technology product of interest to the general population with the potential for a selective choice process. Websites selling consumer electronics (e.g., Best Buy) describe tablet computers on a range of attributes, some of which are likely to be active (e.g., screen size) and some of which are likely to be inactive (e.g., memory technology type) for some people. We use eight such product attributes, each with two levels, as the study stimuli (see Table 7).

Participants in the study were randomly assigned to two groups: a group that received ϕ-designs customized to their self-stated active–inactive attribute status and a second group that received a no-customization D-design. Regardless of group assignment, all participants first completed a self-stated "does matter"–"does not matter" attribute classification task (see Figure 3). Participants randomly assigned to the no-customization group then completed a conjoint choice task based on the benchmark D-design. Participants randomly assigned to the customization group completed a conjoint choice task that had been customized instantaneously using the self-stated information they provided in that first step, the scripting algorithm described in the previous section, and the prebuilt ϕ-design libraries. Both groups answered a total of 18 binary choice questions (for an example, see Figure 4) and completed the common set of 6 holdout tasks as described previously.

Customized Conjoint Choice Experiment Results

A total of 488 adults drawn from a U.S. sample of an online panel were randomly assigned into the no-customization or design customization condition and completed a Qualtrics-enabled study using a web browser. After data quality checks (e.g., "click #2 for this answer"), there were a total of 468 respondents across the no-customization group (N = 233) and the customization group (N = 235). Of the respondents, 44.4% are between 25 and 34 years old, 61.9% own a tablet computer, 34.4% are married, 42.9% are female, and 26.3% have household income between $50,000 and $100,000. Data from each respondent include their self-stated active–inactive attribute classification (z_i), their design matrix (X_i), their selections in the choice task and holdout choice task (y_i), and their random group assignment. There are no significant differences in the active–inactive attribute classification between the two groups for any attributes, nor are there any differences on these demographics between the two groups (all χ^2 tests n.s.).

We fit two models to the data. We estimated the heterogeneous variable selection model that uses the auxiliary data as previously outlined (M1) and a reference random-effects multinomial logit model (M0, where all τ_{im} are fixed to 1) across both the no-customization and customization design groups. Table 8 reports design performance measures for both the no-customization (D-design) and customization (ϕ-design) groups as well as model performance measures for M0 and M1. Consistent with the findings from the simulation study, the results show that customization using ϕ-designs leads to better prediction of choice behavior.

The rightmost two columns in Table 8 report the results of the heterogeneous variable selection model that uses auxiliary data (M1) for the two design strategy groups. Across measures, a comparison of these results shows the superiority of a customized design strategy using the ϕ-criterion. For in-sample performance, the model better fits the data from the customized ϕ-designs.[7] More importantly, better in-sample performance for the customized ϕ-designs also translates into significantly better out-of-sample prediction. The customized ϕ-designs have a 2.8% higher out-of-sample hit rate than the no-customization D-designs.

Performance of the two design strategies under a reference random-effects multinomial logit model (M0; also reported in

Table 8
MODEL AND DESIGN PERFORMANCE MEASURES FROM EMPIRICAL STUDY

	Random-Effects Multinomial Logit (M0)		*Heterogeneous Variable Selection with Auxiliary Data (M1)*	
Model Type	*D-Design*	*ϕ-Design*	*D-Design*	*ϕ-Design*
In-Sample				
DIC	2,498.398	2,277.022	2,394.632	2,109.319
Mean(DIC_i)	6.981	6.267[a]	6.997	6.058[a]
Out-of-Sample				
Hit rate	.698	.721[a]	.709	.738[a]
MAD	.323	.302[a]	.318	.290[a]

[a]95% highest posterior density interval (HPD) of difference between design groups does not include zero.

[7]Because the two groups are not of identical size ($N_{D\text{-}design}$ = 233 vs. $N_{\phi\text{-}design}$ = 235), rather than comparing cumulative group-level deviance information criterion (DIC; Spiegelhalter et al. 2002) values, we compute the group-level mean of the individual-level DIC over the posterior distribution of the estimated model parameters. As with its parent measure, lower values of this measure indicate a better-fitting model.

Table 9
ACTIVE–INACTIVE VARIABLE SELECTION DATA FROM EMPIRICAL STUDY

	Self-Stated Proportion Active	*Fidelity Parameter* (γ)	*Model-Recovered Proportion Active*	*Model-Recovered Inactive Fidelity*	*Model-Recovered Active Fidelity*
	$z_m = 1$		$\tau_m = 1$	$\tau_m = c\|z_m = -1$	$\tau_m = 1\|z_m = 1$
Attribute	*M*	*M (SD)*	*M*	*M*	*M*
Screen size	.953	**.845** (.182)	.688	.682	.677
Storage capacity	.859	**.928** (.231)	.635	.909	.731
Camera resolution	.483	**1.724** (.361)	.474	.967	1.000
Price	.972	**.521** (.157)	.612	.846	.571
Processor type	.632	**1.147** (.214)	.547	.901	.780
Memory technology type	.468	.051 (.344)	.487	.751	.283
Display technology type	.496	.866 (.615)	.483	.992	1.000
Warranty type	.536	.563 (.441)	.498	.894	.928
Across all attributes	.675		.525	.896	.728

Notes: Boldface indicates 95% HPD of mean for fidelity parameter γ does not include zero.

Table 8) provides more insight into the interaction between a priori design customization and model-based accounting for a selective choice process. The reference model (M0) does not formally account for variable selection, but the random-effects specification does include zero as a possible value for an individual parameter. If there are also performance differences between the two design groups under the model without a formal variable selection component (M0), that is strong evidence the design strategy is driving the between-group performance differences. Indeed, the customized ϕ-design group still has significantly better in-sample and out-of-sample performance compared with the no-customization D-design group even when the estimated model does not formally account for variable selection. In contrast, the performance differences across the models within the no-customization D-design group are small. These results suggest that a priori design customization using the proposed ϕ-criterion is more important for making a better inference than is the model itself.

Cost–Benefit Implications of a Customized Design Strategy

In our empirical application, implementing a customized design strategy with the ϕ-criterion requires collecting self-stated data and customizing individual-level designs using that data. The incremental time cost of collecting the "does matter"–"does not matter" attribute classification data in our study for the median respondent was 30.2 seconds, with 80% of respondents taking between 14.2 and 67.6 seconds to complete the task. During data collection, the design customization JavaScript executed on the subsequent page-load event, thus adding no perceptible overhead for study respondents. Overall, our design customization approach requires a prebuilt ϕ-design library encompassing the possible variable selection patterns that may arise in the self-stated data, the assumption of exchangeability across attributes having the same number of levels, and some programming to modify the executable code of the web-based survey. The JavaScript template in the Web Appendix provides an illustration of how this customization can be accomplished in a web-based survey environment.

The advantages of ϕ-designs depend in part on the degree of fidelity between the a priori stated measures about active–inactive attribute status that are used in both design customization as well as model estimation, and the variable selection status inferred from actual choice behavior. In our empirical application, there is a 78.3% concordance between the self-stated data in z and the model-based variable selection parameters τ (summarized in Table 9). Conditioning on the active–inactive classification, there is more concordance ($P(\tau = c|z = -1) = 89.6\%$) between stated inactive and inferred inactive variables (i.e., people stated that the attribute would be inactive and it was indeed inactive as inferred from the actual choices). This suggests that, although there are fewer attributes classified as inactive (32.5%), people are more consistent in their treatment of the inactive attributes between self-stated auxiliary data and the variable selection status inferred from actual choice behavior. Such consistency is fundamental to the aforementioned advantages of ϕ-designs.

In contrast, there is lower concordance ($P(\tau = 1|z = 1) = 72.8\%$) between stated active attributes and active attributes inferred from actual choice behavior. One interpretation of this result is that people are "hedging" when completing the self-stated active–inactive attribute classification task: they might be unwilling to designate an attribute as one that "does not matter" a priori but end up ignoring that attribute in their actual decision making. When designing choice experiments using the ϕ-criterion, this result suggests that the prior distribution for the active attributes should always include the possibility of zero. Although auxiliary data appear to be a more useful proxy for explaining the status of some attributes compared with others, the composite salubrious effect of leveraging the self-stated data in constructing customized choice experiments using the ϕ-criterion and analyzing the consequent data are noteworthy.[8]

As with any ex ante auxiliary data collection task, there is concern that the task itself may alter behavior in a subsequent choice task. Although the two groups in our study both completed the same auxiliary self-stated active–inactive data task (thus, there are no differential treatment concerns), a broader concern is how increased consistency between stated variable selection behavior and variable selection inferred from the choice data because of the proximity of the tasks would

[8]Note that although we have shown that the heterogeneous variable selection model using self-stated auxiliary data performs better than a reference random-effects logit model ($DIC_{M1} = 4{,}503.951$ vs. $DIC_{M0} = 4{,}775.420$), the proposed model also outperforms a heterogeneous variable selection model that does not make use of the auxiliary self-stated data (e.g., Gilbride, Allenby, and Brazell 2006), such that $\theta_m \sim \mathrm{Beta}(a, b)$ (DIC = 4,724.916).

Table 10

PREFERENCE PARAMETERS FOR NO-CUSTOMIZATION D-DESIGN OR CUSTOMIZED ϕ-DESIGN GROUPS FROM THE HETEROGENEOUS VARIABLE SELECTION MODEL WITH AUXILIARY DATA (M1)

	No-Customization D-Design	*Customized ϕ-Design*
Preference Parameter	*M (SD)*	*M (SD)*
Screen size (10.1 inches vs. 7 inches)	**1.604** (.092)	**1.956** (.108)[a]
Storage capacity (32 GB vs. 16 GB)	**1.062** (.062)	**.992** (.063)
Camera resolution (8 megapixel vs. 5 megapixel)	**.407** (.041)	**.433** (.043)
Price ($449 vs. $349)	**−1.468** (.086)	**−1.634** (.092)[a]
Processor type (Quad-core vs. Dual-core)	**.958** (.063)	**.719** (.055)[a]
Memory technology type (LPDDR vs. DDR2)	.080 (.036)	.068 (.035)
Display technology type (IPS LCD vs. TFT LCD)	−.032 (.032)	−.003 (.034)
Warranty type (extended vs. standard)	**.166** (.038)	**.121** (.038)

[a]95% HPD of group mean difference does not include zero.

Notes: Group-level means of individual parameters over the posterior are reported, which also take into account the value of the variable selection parameter at that draw of the posterior distribution. Boldface indicates that the 95% HPD of group mean does not include zero.

favor the proposed ϕ-design approach. Our previous exploration of sparsity misspecification (see Figure 2) demonstrates that the benefits of using an informative prior ϕ-design approach accrue when the probability of stated–inferred concordance is better than chance. More broadly, there is a trade-off between the benefits of using an informative prior and the costs generated in the collection of that data. The use of pilot studies or a time-delay in data collection may mitigate such costs.

An important source of the efficiency and predictive validity advantages of ϕ-designs relative to D-designs is the flexibility afforded in the dual-objective compound design criterion to differentially focus across active–inactive attributes. The upshot of this differential focus on the efficient estimation of the active attributes is more faithful recovery of underlying preferences, as we demonstrated in our simulation studies. The design characteristic differences between the two design strategies suggest that these efficiency advantages should accrue most for the active attributes, which could lead to a different understanding of the marketplace. The results from the heterogeneous variable selection model with auxiliary data (Table 10) show that there are differences in recovered preferences between the customized ϕ-design and no-customization D-design design groups. For example, mean preference for screen size is significantly larger in the customized ϕ-design group compared with the no-customization D-design group ($\beta_{\phi\text{−design}} = 1.965$ vs. $\beta_{\text{D−design}} = 1.604$, prob > .95).[9] With the exception of the design strategy (ϕ- vs. D-design), these differences are not due to other systematic group differences. Recall that the respondents were randomly assigned and, as stated previously, there are no group-level differences in the self-stated auxiliary or demographic data. Moreover, both groups completed the same self-stated auxiliary data collection task, and model estimation results show similar concordance between the auxiliary active–inactive data and the model-inferred variable selection parameters ($P(z{=}\tau)_{\phi\text{−design}} = 78.5\%$ vs. $P(z{=}\tau)_{\text{D−design}} = 78.1\%$).[10]

To illustrate the implications of this result, consider the scenario of designing new products for which the relative value of one attribute versus another is an important managerial consideration. Linear combinations of recovered preference parameters have been used to understand this trade-off (Toubia and Hauser 2007). In our empirical setting, because all attributes are binary, we regard the difference in preference between two attributes as the relative valuation of one attribute versus another. For example, the difference between the mean preference for a larger screen size (10.1 inches vs. 7 inches) and the mean preference for larger storage (32 GB vs. 16 GB) is .964 for the ϕ-design group but only .541 for the D-design group (prob > .95). Screen size is an important attribute for many people (95.3% of respondents designated it as an active attribute). For an analyst using a D-design, these results suggest that the mean trade-off for a larger screen in memory capacity terms is $(.541/1.062) \times 16 \text{ GB} = 8.155$ additional gigabytes, whereas if the analyst used a ϕ-design, the value would be $(.964/.992) \times 16 \text{ GB} = 15.553$ GB. In terms of incremental gigabytes of memory, this difference suggests that an analyst using D-designs could significantly undervalue the larger screen size.[11] If the marginal costs of additional memory and larger screens are different, there could be profitability consequences to decision making based on data obtained from a D-design or a ϕ-design.[12]

Overall, we have shown that the flexibility of the dual-objective ϕ-criterion to allocate differential emphasis across active–inactive attributes leads to more efficient estimation of attribute preferences and, thus, a better understanding of the relative value of those attributes. Although our simulation results support this claim, it is difficult to assert the same in an empirical study. As such, the substantive benefits of a customized ϕ-design strategy can be best summarized in the

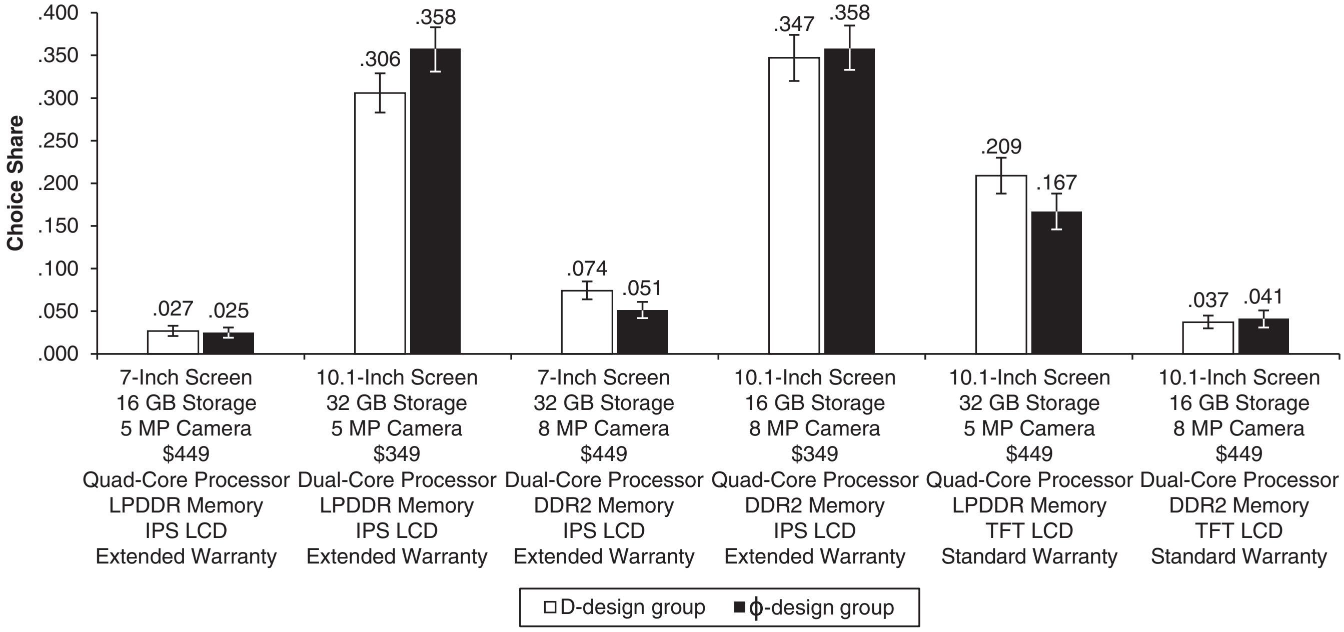
[9]Prob > .95 means that there is >95% probability that the posterior distribution of the difference does not include 0.

[10]A comparison of the posterior variance of preference also reveals no significant differences between the two groups for any of the eight preference parameters. Thus, although some means are different across the groups, preference heterogeneity is not.

[11]Ceteris paribus, an analyst using ϕ-designs would conclude that the average respondent would be indifferent between a 16 GB, 10.1-inch tablet and a ~48 GB, 7-inch tablet, whereas an analyst using D-designs would conclude the average respondent would be indifferent between a 16 GB, 10.1-inch tablet and a ~40 GB, 7-inch tablet.

[12]Although our approach to variable selection can be extended to multilevel attributes, capturing higher-order interactions across attributes or nonlinearities across levels within an attribute may also be of substantive interest; however, these are beyond the scope of the current design and analysis framework.

Figure 5
CHOICE SHARE AMONG SIX PRODUCT ALTERNATIVES FOR D-DESIGN AND ϕ-DESIGN GROUPS



Notes: Each bar represents the D- or ϕ-design group-level average choice share in the depicted six-alternative choice scenario, computed using individual-level parameter estimates over the posterior distribution of those estimates. The error bars represent posterior uncertainty of those means; thus, group-level differences exist for predicted choice shares in which the error bars do not overlap.

superior holdout predictions shown in Table 8. As further illustration, we also presented a choice scenario (described in Figure 5) to the study respondents as a final choice task. In that scenario, there are significant differences between the D-design and ϕ-design groups for the predicted choice shares of products two, three, and five. An analyst using a ϕ-design would conclude that a lower-priced ($349) and less advanced processor type (dual core) tablet computer alternative would garner significantly more share (5.2%; prob > .95), whereas the pricier ($449) and more advanced processor type (quad core) tablet computer would capture significantly less share (4.2%; prob > .95) compared with the results obtained using a D-design. The ϕ-design group has a significantly lower MAD of prediction (.516 vs. .553; prob > .95) than the D-design group for this additional choice question, providing further support to the claimed advantages of the ϕ-designs when it is likely that people will regard some attributes as inactive in their decision making. Because people may differ in both the scope and focus of their own selective choice process, accounting for this possibility in the design and analysis of choice experiments improves predictive performance and has implications for marketing action.

SUMMARY AND DISCUSSION

People often attend to only a subset of active attributes when making choices while ignoring the remaining inactive attributes. In this article, our approach to design and analysis for a selective choice process leverages self-stated auxiliary data as prior information both in individual-level customized design construction and in a heterogeneous variable selection model. The upshot of this novel approach to design and analysis is more efficient experimental design and better model-based inference.

With regard to design, we propose a dual-objective compound design criterion that incorporates prior information to customize an experimental design for each respondent on the basis of self-stated information. This serves the joint purpose of efficiently estimating the effects of active attributes and detecting the effects of attributes that respondents state are inactive but that may turn out to be active. By leveraging self-stated auxiliary data from each respondent as to which attributes are likely to be active in his or her selective choice process, the proposed ϕ-criterion allows for customized designs.[13] In contrast to the traditional D-criterion that focuses on the single objective of efficient estimation of all parameters in the model, the proposed ϕ-criterion is a more flexible design criterion that allows for differential emphasis on variables based on prior information. Results from our comparative study provide strong evidence in favor of using the proposed dual-objective ϕ-criterion rather than the traditional D-criterion when constructing efficient designs for a selective choice process. More broadly, the multiple-objective compound criterion concept we introduce could also be used in other design applications to balance distinct objectives such as efficiency and prediction.

[13]Our approach complements but is distinct from adaptive self-explicated methods (e.g., Netzer and Srinivasan 2011) as well as adaptive conjoint design approaches (e.g., Toubia et al. 2003) that use choice set responses to construct subsequent choice set elements of a design, rather than using auxiliary self-stated data to a priori customize the individual-specific choice design as we do in our approach.

Our integrative approach uses auxiliary data as prior information not only in the design construction through the compound design criterion but also in the analysis through a heterogeneous variable selection model that incorporates respondent-specific auxiliary information about the active–inactive status of an attribute. The simulation results show that, when the choice process is selective, the proposed approach to design and analysis more faithfully recovers individual preferences, has superior in-sample fit, and provides better out-of-sample market predictions. Results from the empirical conjoint choice experiment demonstrate the feasibility of an instantaneous customization strategy using the proposed design criterion and highlight the substantive differences that can arise between the proposed approach and the traditional approach. The results provide strong evidence in favor of using ϕ-designs when it is likely that people will regard some of the attributes as inactive in their decision process.

The costs of implementing individual-specific customized ϕ-designs include the requirement to collect auxiliary data, the need for an instantaneous customization algorithm using a prebuilt design library, and a reasonable expectation that self-stated data on the active–inactive status of attributes is concordant to variable selection behavior inferred from actual choices. The motivation to collect auxiliary active–inactive data is conditioned on the expectation that it will improve both the design and the analysis. The empirical results suggest that there is indeed a useful connection between self-stated data and inferred variable selection status, all attained at a modest respondent time cost.

The current research has focused on individual-level self-stated auxiliary data, individual-specific design customization, and individual inference. When efficient estimation of population-level parameters is of interest, such as the proportion of respondents who attend to (or ignore) an attribute or the mean utility of an active attribute among respondents, the proposed approach may not work well, because designs optimized for the estimation of individual-level parameters may not be optimal for the estimation of population-level parameters (Liu, Dean, and Allenby 2012). A different design criterion would be necessary to find designs optimized for the estimation of population-level parameters (e.g., Liu and Arora 2011). That different criterion would be an aggregate measure that integrates over all possible combinations of variable selection behavior in the target population, where the prior probabilities for variable selection are based on the proportion of respondents who state an attribute to be inactive. This can be computationally prohibitive, especially when there are many attributes and attribute levels. For example, in the case of eight binary attributes with at least one active attribute, this would involve integrating more than $2^8 - 1 = 255$ possibilities. In contrast, the proposed ϕ-criterion avoids such computation as well as the need for ex ante summary statistics about variable selection behavior in the target population by instead utilizing an individual respondent's active–inactive attribute status to customize the design construction for each person in the course of a single data-collection event. In addition, evidence from prior research has suggested that individual-level customization also leads to reasonable efficiency in the estimation of the population-level parameters (Yu, Goos, and Vanderbroek 2011).

In closing, the distinction between active attributes that "do matter" and inactive attributes that "do not matter" in a selective choice process is a unique context that has received limited attention in the marketing literature. Leveraging auxiliary self-stated information in constructing customized designs using the proposed ϕ-criterion allows for the flexibility to differentially focus across active–inactive attributes and provides the capacity to better distinguish between individuals regarding an attribute's status in decision making, leading to a better understanding of individual preferences and more accurate market predictions. When choice experiments are used to inform product decisions in complex, infrequently purchased, or rapidly evolving product categories, failure to account for a selective choice process can have material consequences for the business. The efficiency, accuracy, and predictive validity benefits that a ϕ-design provides in such contexts merits its consideration.

REFERENCES

Atkinson, Anthony C., Alexander N. Donev, and Randall D. Tobias (2007), *Optimum Experimental Designs, with SAS*. Oxford, UK: Oxford University Press.

Aribarg, Anocha, Neeraj Arora, and Moon Young Kang (2010), "Predicting Joint Choice Using Individual Data," *Marketing Science*, 29 (1), 139–57.

Arora, Neeraj, Ty Henderson, and Qing Liu (2011), "Noncompensatory Dyadic Choices," *Marketing Science*, 30 (6), 1028–47.

Arora, Neeraj, and Joel Huber (2001), "Improving Parameter Estimates and Model Prediction by Aggregate Customization in Choice Experiments," *Journal of Consumer Research*, 28 (2), 273–83.

Ben-Akiva, Moshe E., and Steven R. Lerman (1985), *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge, MA: MIT Press.

Bettman, James R., Mary Frances Luce, and John W. Payne (1998), "Constructive Consumer Choice Processes," *Journal of Consumer Research*, 25 (3), 187–217.

Chaloner, Kathryn, and Isabella Verdinelli (1995), "Bayesian Experimental Design: A Review," *Statistical Science*, 10 (3), 273–304.

George, Edward I., and Robert E. McCulloch (1993), "Variable Selection via Gibbs Sampling," *Journal of the American Statistical Association*, 88 (423), 881–89.

Gilbride, Timothy J., and Greg M. Allenby (2004), "A Choice Model with Conjunctive, Disjunctive, and Compensatory Screening Rules," *Marketing Science*, 23 (3), 391–406.

Gilbride, Timothy J., Greg M. Allenby, and Jeff D. Brazell (2006), "Models for Heterogeneous Variable Selection," *Journal of Marketing Research*, 43 (August), 420–30.

Green, Paul E., and Abba M. Krieger (1996), "Individualized Hybrid Models for Conjoint Analysis," *Management Science*, 42 (6), 850–67.

Hensher, David A., and John M. Rose (2009), "Simplifying Choice Through Attribute Preservation or Non-Attendance: Implications for Willingness to Pay," *Transportation Research Part E, Logistics and Transportation Review*, 45 (4), 583–90.

Hensher, David A., John M. Rose, and William H. Greene (2012), "Inferring Attribute Non-Attendance from Stated Choice Data: Implications for Willingness to Pay Estimates and a Warning for Stated Choice Experiment Design," *Transportation*, 39 (2), 235–45.

Huber, Joel, and Klaus Zwerina (1996), "The Importance of Utility Balance in Efficient Choice Designs," *Journal of Marketing Research*, 33 (August), 307–17.

Kessels, Roselinde, Peter Goos, and Martina Vandebroek (2006), "A Comparison of Criteria to Design Efficient Choice Experiments," *Journal of Marketing Research*, 43 (August), 409–19.

Kessels, Roselinde, Bradley Jones, and Peter Goos (2011), "Bayesian Optimal Designs for Discrete Choice Experiments with Partial Profiles," *Journal of Choice Modelling*, 4 (3), 52–74.

Kuhfeld, Warren F., Randall D. Tobias, and Mark Garratt (1994), "Efficient Experimental Design with Marketing Research Applications," *Journal of Marketing Research*, 31 (November), 545–57.

Lenk, Peter, and Bryan Orme (2009), "The Value of Informative Priors in Bayesian Inference with Sparse Data," *Journal of Marketing Research*, 46 (December), 832–45.

Liu, Qing, and Neeraj Arora (2011), "Efficient Choice Designs for a Consider-Then-Choose Model," *Marketing Science*, 30 (2), 321–38.

Liu, Qing, Angela M. Dean, and Greg M. Allenby (2012), "Bayesian Designs for Hierarchical Linear Models," *Statistica Sinica*, 22 (1), 393–417.

Netzer, Oded, and V. Srinivasan (2011), "Adaptive Self-Explication of Multiattribute Preferences," *Journal of Marketing Research*, 48 (February), 140–56.

Sándor, Zsolt, and Michel Wedel (2001), "Designing Conjoint Choice Experiments Using Managers' Prior Beliefs," *Journal of Marketing Research*, 38 (November), 430–44.

Sándor, Zsolt, and Michel Wedel (2002), "Profile Construction in Experimental Choice Designs for Mixed Logit Models," *Marketing Science*, 21 (4), 455–75.

Sándor, Zsolt, and Michel Wedel (2005), "Heterogeneous Conjoint Choice Designs," *Journal of Marketing Research*, 42 (May), 210–18.

Spiegelhalter, David J., Nicola G. Best, Bradley P. Carlin, and Angelika van der Linde (2002), "Bayesian Measures of Model Complexity and Fit," *Journal of the Royal Statistical Society. Series B. Methodological*, 64 (4), 583–639.

Tibshirani, Robert (1996), "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society. Series B. Methodological*, 58 (3), 267–88.

Toubia, Olivier, and John R. Hauser (2007), "On Managerially Efficient Experimental Designs," *Marketing Science*, 26 (6), 851–58.

Toubia, Olivier, Duncan I. Simester, John R. Hauser, and Ely Dahan (2003), "Fast Polyhedral Adaptive Conjoint Estimation," *Marketing Science*, 22 (3), 273–303.

Van Nierop, Erjen, Bart Bronnenberg, Richard Paap, Michel Wedel, and Philip Hans Franses (2010), "Retrieving Unobserved Consideration Sets from Household Panel Data," *Journal of Marketing Research*, 47 (February), 63–74.

Yee, Michael, Ely Dahan, John R. Hauser, and James Orlin (2007), "Greedoid-Based Noncompensatory Inference," *Marketing Science*, 26 (4), 532–49.

Yu, Jie, Peter Goos, and Martina Vanderbroek (2009), "Efficient Conjoint Choice Designs in the Presence of Respondent Heterogeneity," *Marketing Science*, 28 (1), 122–35.

Yu, Jie, Peter Goos, and Martina Vandebroek (2011), "Individually Adapted Sequential Bayesian Conjoint-Choice Designs in the Presence of Consumer Heterogeneity," *International Journal of Research in Marketing*, 28 (4), 378–88.

Zou, Hui, and Trevor Hastie (2005), "Regularization and Variable Selection via the Elastic Net," *Journal of the Royal Statistical Society. Series B. Methodological*, 67 (2), 301–20.