ROSELINDE KESSELS, PETER GOOS, and MARTINA VANDEBROEK*

To date, no attempt has been made to design efficient choice experiments by means of the G- and V-optimality criteria. These criteria are known to make precise response predictions, which is exactly what choice experiments aim to do. In this article, the authors elaborate on the G- and V-optimality criteria for the multinomial logit model and compare their prediction performances with those of the D- and A-optimality criteria. They make use of Bayesian design methods that integrate the optimality criteria over a prior distribution of likely parameter values. They employ a modified Fedorov algorithm to generate the optimal choice designs. They also discuss other aspects of the designs, such as level overlap, utility balance, estimation performance, and computational effectiveness.

A Comparison of Criteria to Design Efficient Choice Experiments

Since Louviere and Woodworth's (1983) article, choice experiments have become increasingly popular to explore consumer preferences for the attributes of various goods. In applied research, these experiments have been used extensively, and in fundamental research, they have been the subject of rigorous study and research. The reason for their popularity is that they enable researchers to model real marketplace choices and thus to emulate real market decisions and predict market demand (Carson et al. 1994). In a typical choice experiment, respondents are presented with a series of choice sets, each composed of several alternatives, also called profiles, of products or services that are defined as combinations of different attribute levels. Respondents then indicate their preferred alternative for every choice set.

Louviere, Street, and Burgess (2003) present an overview of the recent developments in choice experiments, with a special emphasis on the design of these experiments. Designing an efficient choice experiment involves selecting alternatives that, when put into choice sets, provide maximum information on the parameters of a probabilistic choice model. Until now, the efficiency of a choice design has been expressed primarily in terms of the D-optimality criterion (Atkinson and Donev 1992). Only Street, Bunch, and Moore (2001) apply the A-optimality criterion to the design of paired comparison experiments with two-level attributes. To date, the G- and V-optimality criteria, specifically developed for making precise response predictions, have not been applied in the experimental choice context. However, choice experiments are conducted for predictive purposes, and therefore we turn attention to the G- and V-optimality criteria.

The main difficulty in the construction of a proper choice design is that the probabilistic choice models are nonlinear in the parameters, implying that the efficiency of the design depends on the unknown parameter vector (Atkinson and Haines 1996). Consequently, researchers need to assume values for the parameters before deriving the experimental design. To circumvent this circular problem, three approaches have been introduced. We discuss them for logit choice models, the best known of which is the multinomial logit model (McFadden 1974).

The first approach is to use zero prior parameter values so that methods of linear experimental design can be applied. It is implicitly assumed that the respondents prefer all attribute levels and, thus, all alternatives equally (Grossmann, Holling, and Schwabe 2002). The following authors are representatives of this approach: Anderson and Wiley (1992) and Lazari and Anderson (1994) provide a catalog of orthogonal arrays for logit choice models. To address a broader range of design classes, Kuhfeld, Tobias, and Garratt (1994) make use of Cook and Nachtsheim's (1980)

*Roselinde Kessels is a doctoral student, Faculty of Economics and Applied Economics (e-mail: roselinde.kessels@econ.kuleuven.be), and Martina Vandebroek is Professor of Statistics, Faculty of Economics and Applied Economics, and University Center for Statistics (e-mail: martina.vandebroek@econ.kuleuven.be), Katholieke Universiteit Leuven. Peter Goos is Professor of Statistics, Faculty of Applied Economics, Universiteit Antwerpen (e-mail: peter.goos@ua.ac.be). The authors thank the two anonymous *JMR* reviewers for their constructive comments, which led to a substantial improvement of the article. They are also grateful to the Information Technology department of the Faculty of Economics and Applied Economics at the Katholieke Universiteit Leuven for putting some computer rooms at their disposal. They also thank Professor Yuri Goegebeur for his help with the graphical representations in S-PLUS 6.

© 2006, American Marketing Association
ISSN: 0022-2437 (print), 1547-7193 (electronic)

modification of Fedorov's (1972) exchange algorithm to generate D-optimal designs. Kuhfeld and Tobias (2005) continue this line of research by integrating the modified Fedorov algorithm in a comprehensive algorithm, contained in the SAS %MktEx macro, which also exploits a large catalog of orthogonal arrays and Meyer and Nachtsheim's (1995) coordinate-exchange algorithm. Furthermore, Bunch, Louviere, and Anderson (1996) develop the so-called D-optimal shifted or cyclic designs characterized by the minimal level overlap property. This property is satisfied when the frequencies of the attribute levels within a choice set are distributed as equally as possible. The results for D-optimal paired comparison designs for two-level attributes have been described in the work of Street, Bunch, and Moore (2001) and Street and Burgess (2004) and references therein. As we mentioned previously, Street, Bunch, and Moore also compute A-optimal paired comparison designs for two-level attributes. Finally, Burgess and Street (2003) derive D-optimal choice designs for two-level attributes of any choice set size and extend these results in Burgess and Street (2005) to apply to attributes with any number of levels.

The second approach, attributed to the work of Huber and Zwerina (1996), advocates the use of nonzero prior values rather than zero values. The resulting locally D_P-optimal designs prove to be more efficient than the D-optimal designs based on zero prior values. Carlsson and Martinsson (2003) confirm this finding with a comparison study in health economics. To generate the D_P-optimal designs, Huber and Zwerina (1996) propose the relabeling (R) and swapping (S) techniques, shortly referred to as the RS algorithm. The SAS %ChoicEff macro that uses a modified Fedorov algorithm also allows building the designs, as illustrated by Zwerina, Huber, and Kuhfeld (1996; see updated [2005] version).

Finally, the most recent approach has been introduced by Sándor and Wedel (2001) and consists of integrating the associated uncertainty on the assumed parameter values by the use of Bayesian design techniques (Chaloner and Verdinelli 1995). If there is substantial uncertainty about the unknown parameters, the so-called Bayesian D_B-optimal designs outperform the locally D_P-optimal designs. The algorithm used is the RS algorithm and an additional cycling (C) procedure, accordingly called the RSC algorithm. Sándor and Wedel (2002) develop an updated version of this algorithm.

The foregoing researchers have proposed designs for the multinomial logit model to be administered to various respondents whose choices are pooled. As a result, homogeneous parameters across respondents are assumed. In this article, we adopt the same experimental choice scenario to compare the performances of the D-, A-, G-, and V-optimality criteria. Note that we study main effects choice designs only. Our approach is similar to that of Sándor and Wedel (2001) in that we also implement Bayesian design methods. However, we do not apply the RSC algorithm but rather the modified Fedorov algorithm to generate the optimal designs.

The outline of the remainder of the article is as follows: In the next section, we discuss the D-, A-, G-, and V-optimality criteria for the multinomial logit model. After that, we describe the approach to generate the optimal designs with the modified Fedorov algorithm. Then, we present and compare the different optimal designs. The last section concludes the article.

DESIGN CRITERIA FOR THE MULTINOMIAL LOGIT

We depart from a stacked choice design matrix $\mathbf{X} = [\mathbf{x}'_{js}]_{j=1,\ldots,J;\, s=1,\ldots,S}$, where $\mathbf{x}_{js}$ is a k × 1 vector of the attribute levels of profile j in choice set s. The utility a person attaches to that profile is modeled by $U_{js} = \mathbf{x}'_{js}\boldsymbol{\beta} + \varepsilon_{js}$, where $\boldsymbol{\beta}$ is a k × 1 vector of parameters and ε_{js} is an i.i.d. extreme value error term. With these notations, the multinomial logit probability that profile j in choice set s is chosen amounts to $p_{js} = e^{\mathbf{x}'_{js}\boldsymbol{\beta}}/\Sigma_{t=1}^{J} e^{\mathbf{x}'_{ts}\boldsymbol{\beta}}$. The information matrix, which is the inverse of the variance–covariance matrix of the parameter estimators, is obtained as

$$\mathbf{I}(\mathbf{X},\boldsymbol{\beta}) = N\sum_{s=1}^{S}\mathbf{X}'_s(\mathbf{P}_s - \mathbf{p}_s\mathbf{p}'_s)\mathbf{X}_s, \tag{1}$$

where $\mathbf{X}_s = [\mathbf{x}_{1s}, \ldots, \mathbf{x}_{Js}]'$, $\mathbf{p}_s = [p_{1s}, \ldots, p_{Js}]'$, $\mathbf{P}_s = \text{diag}[p_{1s}, \ldots, p_{Js}]$, and N is the number of respondents. In optimal design theory (Atkinson and Donev 1992; Fedorov 1972; Silvey 1980), direct functions of the information matrix, referred to as optimality criteria or design criteria, are proposed to generate optimal designs that yield precise parameter estimates or accurate predictions. However, the information matrix in Equation 1 depends on the unknown parameters through the probabilities so that parameter values are required before optimal choice designs can be constructed. As we mentioned previously, Sándor and Wedel (2001) adopt a Bayesian design approach that involves the specification of a prior parameter distribution $\pi(\boldsymbol{\beta})$. Usually, this distribution is the normal distribution, $N(\boldsymbol{\beta}|\boldsymbol{\beta}_0,\boldsymbol{\Sigma}_0)$, from which R prior parameters $\boldsymbol{\beta}^r$, r = 1, ..., R, are drawn to approximate it. In general, the resulting Bayesian optimal designs outperform locally optimal designs that are based on a single prior parameter.

D- and A-Optimality Criteria

The most popular optimality criterion to design choice experiments is the D-optimality criterion. The D-optimality criterion seeks to maximize the determinant of the information matrix in Equation 1, or to minimize its inverse, the determinant of the variance–covariance matrix of the parameter estimators. It is related to the A-optimality criterion, which prefers the design for which the sum or the average of the variances of the parameter estimators is minimized. However, a drawback of the A-optimality criterion is that the ordering of designs with respect to this criterion depends on the type of coding. We refer to the work of Goos (2002, pp. 38–40) for an example in the case of linear models. Note that the A-optimality criterion is more suited for obtaining precise parameter estimates because it considers the variances of the estimators only. In addition, the D-optimality criterion takes the covariances into account. Formally, the Bayesian D-criterion value, the D_B-criterion value, is

$$D_B = \int_{\mathbf{R}^k}\left\{\det\left(\mathbf{I}^{-1}(\mathbf{X},\boldsymbol{\beta})\right)\right\}^{1/k}\pi(\boldsymbol{\beta})d\boldsymbol{\beta}, \tag{2}$$

where the exponent 1/k ensures that it is independent of the dimension k of the parameter vector $\boldsymbol{\beta}$. Minimizing this value results in the D_B-optimal design. The A_B-optimal design minimizes to

$$A_B = \int_{\mathbf{R}^k} \operatorname{tr}\left(\mathbf{I}^{-1}(\mathbf{X},\boldsymbol{\beta})\right)\pi(\boldsymbol{\beta})d\boldsymbol{\beta}. \tag{3}$$

G- and V-Optimality Criteria

The G- and V-optimality criteria both look for designs that make precise predictions about the response. Although choice experiments are carried out to make precise predictions about consumers' future purchasing behavior, the G- and V-optimality criteria have not yet been applied in the experimental choice context. To derive the G- and V-optimality criteria for the nonlinear choice model, predicted probabilities must be computed, and to do so, choice sets must be specified. In particular, we computed the predicted probabilities with respect to all possible choice sets of size J that can be composed from the candidate profiles. These choice sets make up the design region χ. Thus, if there are Q possible choice sets, $\chi = \{\{\mathbf{x}_{1q}, ..., \mathbf{x}_{Jq}\}|q = 1, ..., Q\}$. Then, by definition, the G-optimality criterion aims to minimize the maximum prediction variance over the design region χ, whereas the V-optimality criterion aims to minimize the average prediction variance over this region. Mathematically, the G_B-criterion value is given by

$$\begin{aligned} G_B &= \int_{\mathbf{R}^k} \max_{\mathbf{x}_{jq}\in\chi} \operatorname{var}\left\{\hat{p}_{jq}\left(\mathbf{x}_{jq},\boldsymbol{\beta}\right)\right\}\pi(\boldsymbol{\beta})d\boldsymbol{\beta} \\ &= \int_{\mathbf{R}^k} \max_{\mathbf{x}_{jq}\in\chi} \mathbf{c}'\left(\mathbf{x}_{jq}\right)\mathbf{I}^{-1}(\mathbf{X},\boldsymbol{\beta})\mathbf{c}\left(\mathbf{x}_{jq}\right)\pi(\boldsymbol{\beta})d\boldsymbol{\beta}, \end{aligned} \tag{4}$$

where $\hat{p}_{jq}(\mathbf{x}_{jq},\boldsymbol{\beta})$ denotes the predicted choice probability for $\mathbf{x}_{jq}$ and

$$\mathbf{c}\left(\mathbf{x}_{jq}\right) = \frac{\partial p_{jq}\left(\mathbf{x}_{jq},\boldsymbol{\beta}\right)}{\partial\boldsymbol{\beta}} \tag{5}$$

is the first-order truncated Taylor series expansion of the multinomial logit probability. This approach is similar to the computation of locally D- and c-optimal designs for nonlinear models in general (Atkinson and Donev 1992; Atkinson and Haines 1996). Using the multinomial logit model, we can write Equation 5 as

$$\mathbf{c}\left(\mathbf{x}_{jq}\right) = p_{jq}\left(\mathbf{x}_{jq} - \sum_{t=1}^{J} p_{tq}\mathbf{x}_{tq}\right). \tag{6}$$

Akin to the A-optimality criterion, the relative design efficiency in terms of the G-optimality criterion is contingent on the type of coding. The same applies to the V-optimality criterion, the Bayesian value of which is obtained as

$$V_B = \int_{\mathbf{R}^k}\int_{\chi} \mathbf{c}'\left(\mathbf{x}_{jq}\right)\mathbf{I}^{-1}(\mathbf{X},\boldsymbol{\beta})\mathbf{c}\left(\mathbf{x}_{jq}\right)d\mathbf{x}_{jq}\pi(\boldsymbol{\beta})d\boldsymbol{\beta}, \tag{7}$$

with $\mathbf{c}(\mathbf{x}_{jq})$ defined in Equations 5 and 6.

In Appendix A, we present a simple numerical example of the construction of Bayesian optimal designs by means of the D_B-, A_B-, G_B-, and V_B-optimality criteria.

ALGORITHMIC APPROACH

The most embedded algorithms in the literature to generate choice designs are the RSC algorithm, embracing relabeling, swapping, and cycling, and the modified Fedorov algorithm. There are two versions of the RSC algorithm developed by Sándor and Wedel, one in 2001 and one in 2002. As opposed to the first version, the updated version does not restrict its searches to designs that satisfy the minimal level overlap property, provided that the starting design complies with it. This makes the RSC algorithm more prone to design improvements. As a result of its modification, the RSC algorithm generates designs that are statistically as efficient as those produced by the modified Fedorov algorithm. In the modified Fedorov algorithm, design profiles are exchanged with the profiles from a predefined set of candidate profiles without the enforcement of any constraint. We prefer the modified Fedorov algorithm because it is faster than the adjusted RSC algorithm in generating Bayesian optimal designs. Thus, we incorporated the D_B-, A_B-, G_B-, and V_B-optimality criteria in the modified Fedorov algorithm to obtain four distinct Bayesian modified Fedorov choice algorithms. To avoid poor local optima, we repeated each of the algorithms for several starting designs. We refer to each repetition as a "try," and we performed 200 tries.

With the Bayesian modified Fedorov choice algorithms, we constructed D_B-, A_B-, G_B-, and V_B-optimal designs of two classes. The first class is given by designs of type $3^2 \times 2/2/12$—that is, designs with 12 choice sets, each of size two, in which each alternative is described by three attributes. The first two attributes have three levels each, and the third attribute has two levels. The designs in the second class are of type $3^2 \times 2/3/8$, comprising 8 choice sets of size three and a similar attribute structure as the first design class. As a result, the sets of candidate profiles of both design classes are identical, enclosing the same $3^2 \times 2 = 18$ profiles. In addition, the designs of the two classes consist of the same number of profiles (i.e., 24) to compare the two- and three-alternative optimal designs with respect to specific design measures (see the "Results" section). To compute the G_B- and V_B-optimal designs, the design region χ needs to be specified for each class. For the two-alternative design class, χ consists of $Q = \binom{18}{2} = 153$ choice sets, or 306 profiles, whereas for the three-alternative design class, it includes $Q = \binom{18}{3} = 816$ choice sets, or 2448 profiles. Furthermore, through the use of effects-type coding (see Appendix A), the number of parameter values, k, is five. As prior parameter distribution, we used the multivariate normal distribution $\pi(\boldsymbol{\beta}) = N(\boldsymbol{\beta}|\boldsymbol{\beta}_0,\boldsymbol{\Sigma}_0)$, with $\boldsymbol{\beta}_0 = [-1, 0, -1, 0, -1]'$ and $\boldsymbol{\Sigma}_0$ the 5×5 identity matrix. The $\boldsymbol{\beta}_0$ vector is special because the values for the levels of each of the attributes are equally spaced between –1 and 1. Through this scaling, the utilities increase with the levels of each attribute. For example, for the first two attributes that possess three levels each, a utility of –1 is attached to level 1, a utility of 0 to level 2, and a utility of 1 to level 3. A more extensive account on the specification of $\boldsymbol{\beta}_0$ can be found in the work of Huber and Zwerina (1996). Following Sándor and Wedel's (2001) example, we drew R = 1000 samples $\boldsymbol{\beta}^r$ from $\pi(\boldsymbol{\beta})$.

RESULTS

In Part 1 of this section, we compare the two- and three-alternative D_B-, A_B-, G_B-, and V_B-optimal designs with respect to their performances on several measures, and in Part 2, we take some computational aspects into account. We performed all computations with the SAS 8.02 procedure IML (Interactive Matrix Language).

Part 1: Performance of the D_B-, A_B-, G_B- and V_B-Optimality Criteria

We begin by illustrating the two- and three-alternative D_B-, A_B-, G_B-, and V_B-optimal designs, followed by a study of their amount of level overlap and degree of utility balance. We then score the robustness of the designs on other design criteria for which they are not optimized. Finally, we discuss the accuracy and predictive validity of the parameter estimates.

Designs, overlap, and utility balance. The two- and three-alternative Bayesian optimal designs appear in Tables 1 and 2, respectively. Their criterion values appear in Table 3. The designs clearly exhibit some level overlap. As in the work of Sándor and Wedel (2002), we computed the percentage of the cases in which the columns of the choice sets do not satisfy the minimal level overlap property. The results appear in Table 4. The D_B-optimal designs have the lowest level overlap in the two design classes, followed by the V_B-optimal designs. In contrast, the G_B- and A_B-optimal designs have the highest level overlap.

To measure the utility balance of the computed designs, we built on the cumulative entropy of a choice design, as suggested by Swait and Adamowicz (2001). Utility balance is a concept that Huber and Zwerina (1996) introduced, and it refers to the situation in which respondents prefer the alternatives in a choice set equally and thus face a difficult choice task. In the Bayesian framework, the cumulative entropy of a choice design is defined as

$$\text{(8)} \qquad \mathrm{CH}(\mathbf{X},\boldsymbol{\beta}) = -\sum_{s=1}^{S}\int_{\mathbf{R}^k}\left(\sum_{j=1}^{J}\left\{p_{js}(\mathbf{X}_s,\boldsymbol{\beta})\ln\left(p_{js}(\mathbf{X}_s,\boldsymbol{\beta})\right)\right\}\right)\pi(\boldsymbol{\beta})d\boldsymbol{\beta}.$$

To derive the lower and upper bounds for the cumulative entropy for the two design classes, we constructed for each class a Bayesian design that is not utility balanced and a

Table 1

TWO-ALTERNATIVE BAYESIAN OPTIMAL DESIGNS

		D_B			A_B			G_B			V_B		
		Attributes			*Attributes*			*Attributes*			*Attributes*		
Choice Set	*Alternative*	*1*	*2*	*3*	*1*	*2*	*3*	*1*	*2*	*3*	*1*	*2*	*3*
1	1	2	3	1	3	2	1	3	1	2	2	2	2
	2	1	2	1	3	1	1	2	2	2	1	1	1
2	1	2	2	2	2	3	1	3	2	1	2	1	2
	2	1	1	1	1	2	1	2	3	1	1	2	1
3	1	1	2	2	2	1	2	1	2	1	1	2	2
	2	3	1	2	1	2	2	2	1	1	3	1	1
4	1	2	2	1	3	1	1	1	3	1	2	2	1
	2	1	3	1	2	2	1	3	1	2	1	3	2
5	1	2	2	1	2	2	1	3	3	2	2	1	1
	2	3	3	2	1	1	1	2	1	1	3	2	2
6	1	2	1	2	3	3	2	2	3	1	1	3	1
	2	1	2	1	2	1	1	1	1	1	2	1	1
7	1	1	1	2	1	3	1	1	3	1	1	2	1
	2	2	2	2	2	2	1	2	1	1	3	3	2
8	1	1	2	2	1	3	1	2	2	2	3	2	1
	2	2	1	1	3	1	2	1	1	1	2	3	1
9	1	3	2	1	3	2	1	3	2	2	3	2	2
	2	2	1	1	2	3	2	1	2	2	1	1	2
10	1	3	1	1	1	2	2	1	2	1	2	3	1
	2	2	3	2	2	1	1	1	3	2	1	1	1
11	1	1	3	1	1	3	2	3	2	1	2	3	1
	2	3	1	2	3	3	2	2	2	2	2	2	2
12	1	2	1	1	2	1	2	1	2	2	3	1	2
	2	3	2	2	1	1	1	2	2	1	2	2	2

Table 2
THREE-ALTERNATIVE BAYESIAN OPTIMAL DESIGNS

		D_B			A_B			G_B			V_B		
		Attributes			*Attributes*			*Attributes*			*Attributes*		
Choice Set	*Alternative*	*1*	*2*	*3*	*1*	*2*	*3*	*1*	*2*	*3*	*1*	*2*	*3*
1	1	3	2	1	1	3	2	2	2	1	1	2	2
	2	2	1	1	2	3	1	3	1	2	3	1	1
	3	1	2	2	1	2	1	1	3	1	2	2	1
2	1	1	1	1	1	1	1	3	2	2	3	2	1
	2	2	2	1	2	1	1	1	1	2	1	3	1
	3	1	3	2	1	2	1	3	1	2	2	1	1
3	1	1	3	1	3	1	1	2	3	2	2	1	1
	2	2	3	2	2	2	1	1	2	1	1	2	1
	3	2	1	1	2	2	2	2	3	1	1	1	2
4	1	3	1	2	1	1	1	3	3	2	3	3	2
	2	1	2	1	1	3	2	3	1	1	2	2	2
	3	2	1	1	2	2	2	2	2	2	2	1	1
5	1	3	1	1	1	3	2	2	1	1	1	2	1
	2	3	2	2	2	1	2	2	3	2	2	3	1
	3	1	2	1	3	2	2	3	2	2	3	3	2
6	1	2	1	2	2	2	1	1	3	1	2	3	2
	2	1	2	1	3	2	2	2	3	2	3	1	2
	3	2	3	1	1	1	1	2	1	2	1	3	1
7	1	3	2	2	1	2	1	1	2	2	3	3	1
	2	1	1	1	2	1	2	2	1	2	3	2	2
	3	2	3	2	2	2	2	3	1	2	1	3	1
8	1	1	3	1	3	1	1	1	2	2	2	3	2
	2	2	2	1	1	3	1	1	1	1	1	1	1
	3	3	1	1	3	3	1	2	2	1	2	2	1

Bayesian design that is utility balanced. We obtained the former design by selecting the S choice sets that produced the smallest Bayesian cumulative entropy out of all possible ones, whereas we initially constructed the latter by choosing the S choice sets that produced the largest Bayesian cumulative entropy. This could easily be done by enumerating all Q possible choice sets of size J. However, in doing so, the Bayesian utility balanced designs turned out to be singular. We solved this problem by optimally replacing a minimum number of choice sets with choice sets with a slightly smaller Bayesian entropy. For the two-alternative design class, the values for the minimum and maximum Bayesian cumulative entropy are equal to 1.93 and 5.53, respectively, whereas for the three-alternative design class, these values amount to 2.00 and 5.97.

Subsequently, we computed the values of the cumulative entropy for the two- and three-alternative Bayesian optimal designs and compared them with their maximum value. The values and their percentages appear in Table 5. On the whole, the designs are not maximum utility balanced but entail a moderate choice task complexity. This finding is counter to Huber and Zwerina's (1996) statement that proper choice designs must be maximum utility balanced. Although not perfectly utility balanced, the A_B-optimal designs display the largest cumulative entropy, which extends Arora and Huber's (2001) result that A_P-optimal designs for binary logit models are utility balanced to a Bayesian context. Conversely, the V_B-optimal designs, which are developed especially for making precise predictions, exhibit the smallest cumulative entropy or the least complicated choice tasks. In addition, the three-alternative

Table 3
D_B-, A_B-, G_B-, AND V_B-CRITERION VALUES OF THE TWO- AND THREE-ALTERNATIVE D_B-, A_B-, G_B-, AND V_B-OPTIMAL DESIGNS

Optimal	*Number of Alternatives*	
Design	*2*	*3*
D_B	.73024	.76617
A_B	6.60563	6.02261
G_B	.51997	.51843
V_B	.07219	.06285

Table 4
PERCENTAGES OF LEVEL OVERLAP IN THE TWO- AND THREE-ALTERNATIVE BAYESIAN OPTIMAL DESIGNS

Optimal	*Number of Alternatives*	
Design	*2*	*3*
D_B	14%	38%
A_B	28%	63%
G_B	28%	54%
V_B	17%	42%

designs appear to be more complex than the two-alternative ones.

Performance in terms of other optimality criteria. Because the D_B-, A_B-, G_B-, and V_B-optimality criteria all have a different aim, it is interesting to observe how robust the Bayesian optimal designs are to other design criteria for which they are not optimized. Panels A and B of Table 6 give the efficiencies of the two- and three-alternative optimal designs with respect to the different optimality criteria. As we expected from optimal design theory, the efficiencies of the D_B-optimal designs on the A_B-optimality criterion and of the A_B-optimal designs on the D_B-optimality criterion are high. This interdependence of criterion efficiencies also occurs between the G_B- and V_B-optimality criteria. Furthermore, compared with the D_B-optimal designs, the A_B-optimal designs do not score well in terms of G_B- and V_B- efficiency. As a result, the predictive ability of the A_B-optimal designs is relatively low.

Accuracy and predictive validity of the parameter estimates. We now examine more closely the accuracy and predictive validity of the parameter estimates produced by the two- and three-alternative Bayesian optimal designs. To this end, we investigate the expected mean square errors of the parameter estimates, $EMSE_{\hat{\boldsymbol{\beta}}}$, and of the predicted probabilities, $EMSE_{\hat{\mathbf{p}}_c}$. The $EMSE_{\hat{\boldsymbol{\beta}}}$ pertains to the accuracy of the parameter estimates and is given by

$$\text{(9)} \qquad \mathrm{EMSE}_{\hat{\boldsymbol{\beta}}}(\boldsymbol{\beta}) = \int_{\mathbf{R}^k} (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) f(\hat{\boldsymbol{\beta}}) d\hat{\boldsymbol{\beta}},$$

Table 5

VALUES AND PERCENTAGE VALUES OF CUMULATIVE ENTROPY OF THE TWO- AND THREE-ALTERNATIVE BAYESIAN OPTIMAL DESIGNS

Optimal Design	*Number of Alternatives* 2		3	
D_B	3.98	72%	4.41	74%
A_B	4.27	77%	4.70	79%
G_B	3.96	72%	4.55	76%
V_B	3.72	67%	4.29	72%

Table 6

PERFORMANCES OF THE BAYESIAN OPTIMAL DESIGNS IN TERMS OF OTHER DESIGN CRITERIA

Evaluation Criterion	D_B	A_B	G_B	V_B
A: Two-Alternative Designs				
D_B	100.00%	90.82%	89.08%	93.28%
A_B	97.59%	100.00%	92.43%	87.13%
G_B	94.49%	85.68%	100.00%	99.68%
V_B	96.95%	88.12%	96.03%	100.00%
B: Three-Alternative Designs				
D_B	100.00%	93.49%	94.13%	96.36%
A_B	94.03%	100.00%	96.63%	89.80%
G_B	80.81%	80.19%	100.00%	95.04%
V_B	95.65%	86.04%	96.24%	100.00%

where $f(\hat{\boldsymbol{\beta}})$ is the distribution of the estimates. The smaller the $EMSE_{\hat{\boldsymbol{\beta}}}$ value, the more accurately the parameters are estimated. The $EMSE_{\hat{\mathbf{p}}_c}$ concerns the predictions with respect to the design that contains all Q possible choice sets of size J. This design is chosen so as not to favor any optimal design and is referred to as the complete choice design. It contains the same Q choice sets as the design region χ we defined previously. To compare the prediction performances of the two- and three-alternative optimal designs, we averaged the $EMSE_{\hat{\mathbf{p}}_c}$ values over the number of profiles in the complete choice design. Formally,

$$\text{(10)} \qquad \mathrm{EMSE}_{\hat{\mathbf{p}}_c}(\boldsymbol{\beta}) = \frac{1}{J \times Q} \int_{\mathbf{R}^k} \left(\hat{\mathbf{p}}_c(\hat{\boldsymbol{\beta}}) - \mathbf{p}_c(\boldsymbol{\beta})\right)' \left(\hat{\mathbf{p}}_c(\hat{\boldsymbol{\beta}}) - \mathbf{p}_c(\boldsymbol{\beta})\right) f(\hat{\boldsymbol{\beta}}) d\hat{\boldsymbol{\beta}},$$

where $\mathbf{p}_c(\boldsymbol{\beta})$ is the vector of true logit probabilities in the complete choice design and $\hat{\mathbf{p}}_c(\hat{\boldsymbol{\beta}})$ is the corresponding vector of predicted logit probabilities. The smaller the $EMSE_{\hat{\mathbf{p}}_c}$ value, the more precisely the probabilities are predicted. The distribution of the parameter estimates, $f(\hat{\boldsymbol{\beta}})$, in Equations 9 and 10 is approximated by estimating the parameter values 1000 times using simulated choices from N = 50 respondents. Because the EMSE measures depend on a true parameter $\boldsymbol{\beta}$, we repeated their computation 50 times, each time for a different true parameter. Each computation for another $\boldsymbol{\beta}$ is called a "replication."

We summarize the results of the 50 replications of the $EMSE_{\hat{\boldsymbol{\beta}}}$ in Panel A of Table 7 and in Figure 1. Using percentage values, Table 7, Panel A, depicts the number of replications for which the two- and three-alternative D_B-, A_B-, G_B-, and V_B-optimal designs have the lowest $EMSE_{\hat{\boldsymbol{\beta}}}$ value. The values themselves appear in box plots in Figure 1. The white line in each of the boxes is the median. Overall, for the two design classes, it appears that there is no salient optimality criterion that leads to the most accurate estimates. From the box plots, we observe that the estimation performances of the design criteria are comparable. The median $EMSE_{\hat{\boldsymbol{\beta}}}$ values and the average $EMSE_{\hat{\boldsymbol{\beta}}}$ values,

Table 7

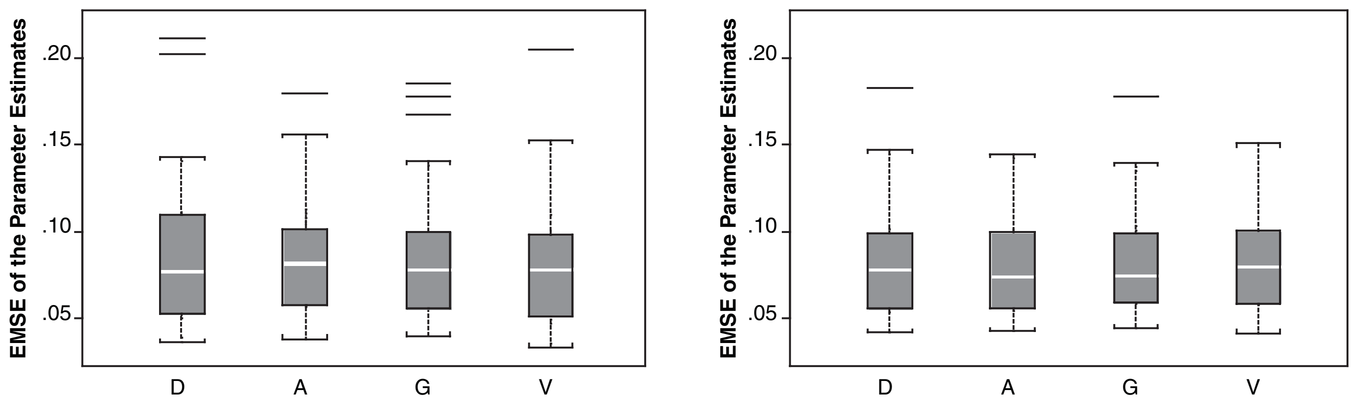
PERCENTAGE OF REPLICATIONS WITH THE LOWEST VALUES FOR THE $EMSE_{\hat{\beta}}$ AND $EMSE_{\hat{p}_c}$ BETWEEN THE TWO- AND THREE-ALTERNATIVE BAYESIAN OPTIMAL DESIGNS

Optimal Design	*Number of Alternatives* 2	3
A: Replications with Lowest $EMSE_{\hat{\beta}}$		
D_B	26%	32%
A_B	16%	28%
G_B	14%	18%
V_B	44%	22%
B: Replications with Lowest $EMSE_{\hat{p}_c}$		
D_B	18%	14%
A_B	6%	10%
G_B	16%	22%
V_B	60%	54%

Figure 1

DISTRIBUTIONS OF THE $EMSE_{\hat{\beta}}$ OBTAINED FROM 50 REPLICATIONS AND COMPUTED FOR THE TWO- AND THREE-ALTERNATIVE D_B-, A_B-, G_B-, AND V_B-OPTIMAL DESIGNS

A: Two-Alternative Designs *B: Three-Alternative Designs*



which are practically identical to the medians but are not shown in the box plots, are equal across the different optimality criteria. Furthermore, Table 7, Panel A, indicates that the G_B-optimality criterion has the smallest number of replications with the lowest $EMSE_{\hat{\beta}}$ value. With regard to estimation differences between the two- and three-alternative designs, the box plots reveal that occasionally, larger $EMSE_{\hat{\beta}}$ values are obtained for the two-alternative designs than for the three-alternative designs. Therefore, the parameter estimates produced by the three-alternative designs are somewhat more accurate.

We carried out an analogous study for the 50 replications of the $EMSE_{\hat{p}_c}$. The number of replications with the lowest $EMSE_{\hat{p}_c}$ value for each of the optimal designs appears in Panel B of Table 7. The box plots with the $EMSE_{\hat{p}_c}$ values for the two- and three-alternative optimal designs appear in Figure 2. As is illustrated by the medians in the plots, the occurrence of larger $EMSE_{\hat{p}_c}$ values for the two-alternative than for the three-alternative designs is more pronounced. As a result, predictions based on the three-alternative designs tend to be more precise. With respect to the prediction performances of the optimality criteria, there are no real surprises. Table 7, Panel B, and the box plots palpably point toward the V_B-optimality criterion as the criterion that provides the most precise predictions. The G_B-optimality criterion is second best, followed by the D_B-optimality criterion and the A_B-optimality criterion. From the box plots, it is also apparent that the predictive capabilities of the customarily used D_B-optimal designs do not differ that much from those of the V_B- and G_B-optimal designs, which are developed particularly for predictive purposes. Therefore, the D_B-optimal designs seem to perform reasonably well in this respect.

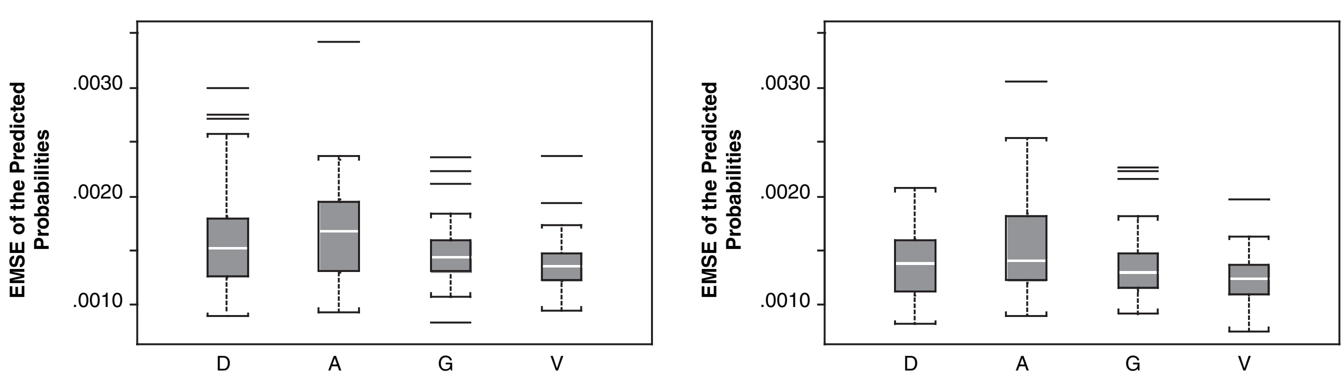
Part 2: Some Computational Aspects

We now embark on the account of some computational aspects of the two- and three-alternative D_B-, A_B-, G_B-, and V_B-optimal designs. We consecutively discuss the computa-

Figure 2

DISTRIBUTIONS OF THE $EMSE_{\hat{p}_c}$ OBTAINED FROM 50 REPLICATIONS AND COMPUTED FOR THE TWO- AND THREE-ALTERNATIVE D_B-, A_B-, G_B-, AND V_B-OPTIMAL DESIGNS

A: Two-Alternative Designs *B: Three-Alternative Designs*



tion times to generate the designs and the computational effectiveness of the design criteria.

Computation time. Table 8 reports computation times for one try of the modified Fedorov algorithm to produce the two- and three-alternative D_B-, A_B-, G_B-, and V_B-optimal designs. We generated the designs with the SAS 8.02 procedure IML. We obtained the times using a Dell personal computer with a 1.80 GHz Intel Processor and 256 MB RAM. Overtly, the computation times for the G_B- and V_B-optimal designs are much longer than those for the D_B- and A_B-optimal designs. This is because of the numerous prediction variances that need to be computed when evaluating a design by means of the G_B- or V_B-optimality criterion. Furthermore, the number of prediction variances to derive is proportional to the design region χ, which is eight times larger for the three-alternative design class than for the two-alternative class. This explains why it takes much more time to construct G_B- and V_B-optimal designs with three alternatives than with two alternatives.

Computational effectiveness of the design criteria. The computational effectiveness of a Bayesian design criterion refers to the quality and the speed of the modified Fedorov algorithm in which this criterion is integrated. We compared the computational effectiveness of the Bayesian design criteria by means of the estimated expected efficiencies from several numbers of tries. The estimated expected efficiency of an optimal design produced by a number of tries, T, is defined as the efficiency to expect when T tries have been performed. We explain the calculation of the expected efficiency from T tries in Appendix B. For each of the two- and three-alternative D_B-, A_B-, G_B-, and V_B-optimal designs, we plotted the expected efficiencies against various numbers of tries. The plots appear in Figure 3. We obtained the highest expected efficiencies when we used the D_B- and A_B-optimality criteria. Applying the G_B- and V_B-optimality criteria requires more tries to reach a given efficiency. Consequently, the smallest number of tries is needed for calculating the D_B- and A_B-optimal designs. In addition, if the algorithm fails to find the D_B- and A_B-optimal designs, it still produces highly efficient designs.

CONCLUSION

In this article, we incorporated the D_B-, A_B-, G_B-, and V_B-optimality criteria in the modified Fedorov algorithm to generate two- and three-alternative Bayesian optimal choice designs containing the same number of profiles. We devoted special attention to the G_B- and V_B-optimality criteria, which look for designs that produce precise predictions. After all, choice experiments are carried out to predict the future market share of related products or services as precisely as possible. We observed that the V_B-optimal designs and, to a lesser extent, the G_B-optimal designs are best suited for predictive purposes. The D_B-optimal designs rank third in this aspect, but the differences in predictive ability compared with the V_B- and G_B-optimal designs are rather small. Furthermore, the three-alternative optimal designs lead to better predictions than the two-alternative designs. The three-alternative optimal designs also yield the most accurate parameter estimates, but there is no real difference in estimation performance between the distinct optimality criteria. However, the computation of the V_B- and G_B-optimal designs takes a long time, particularly those with three alternatives, and many tries are needed. The D_B- and A_B-optimal designs are much faster to compute. To speed up the computations, the number of prior parameter values drawn from the prior distribution can be slightly reduced when evaluating a design. Nevertheless, in weighing the large computational efforts against the small improvements in predictive ability of the V_B- and G_B-optimal designs, it seems preferable to retain the use of the D_B-optimality criterion to build optimal choice designs. Moreover, as a rule of thumb, we cogently argue that if more than three attributes with more than two levels each are involved in the design optimization, the use of the V_B- and G_B-optimality criteria in combination with the modified Fedorov algorithm is no longer practically feasible. Drawing on the V_B- and G_B-optimality criteria to deal with large problem situations awaits the exploration of computationally more efficient algorithms. Finally, the Bayesian optimal designs are characterized by some level overlap and are not maximum utility balanced.

Table 8

COMPUTATION TIMES FOR ONE TRY OF THE MODIFIED FEDOROV ALGORITHM TO GENERATE THE TWO- AND THREE-ALTERNATIVE BAYESIAN OPTIMAL DESIGNS

Design Criterion	*Number of Alternatives* 2	3
D_B	00:05	00:05
A_B	00:05	00:05
G_B	02:30	11:00
V_B	02:30	11:00

Notes: Times are expressed in hours:minutes.

APPENDIX A

We compute the D_B-, A_B-, G_B-, and V_B-criterion values of a small design consisting of three choice sets with two alternatives each. The alternatives include two attributes: Attribute 1 has three levels, and Attribute 2 has two levels. The design matrix can be composed either by assigning numerical values (i.e., 1, 2, 3, and so forth) to the attribute levels or by employing effects-type coding. However, because of the categorical nature of the explanatory variables, it is more common to work with the design matrix from effects-type coding. With effects-type coding, the three levels of Attribute 1 are coded as [1 0], [0 1], and [–1 –1], and the two levels of Attribute 2 are coded as –1 and 1. The design matrix containing numerical values, $\mathbf{X}^0$, and its companion in effects-type coding, $\mathbf{X}$, which we use in the computations, appear as follows:

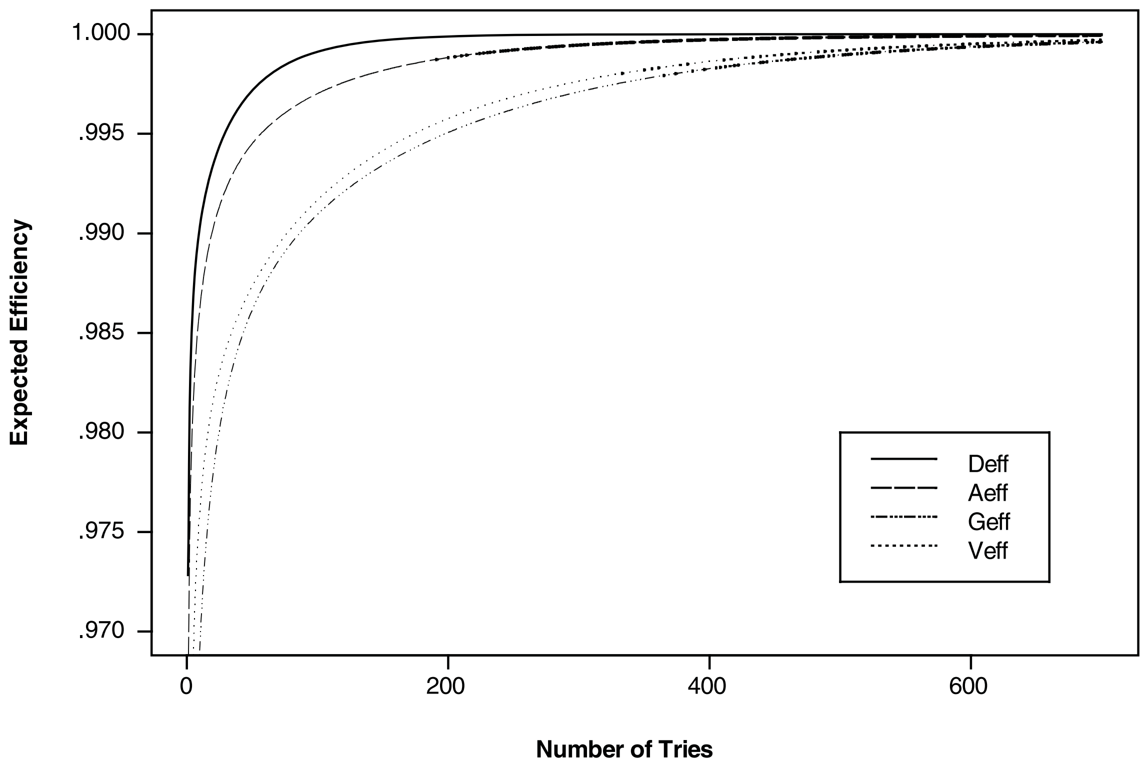
$$\text{(A1)} \qquad \mathbf{X}^0 = \left[\begin{array}{cc} 1 & 2 \\ 2 & 1 \\ \hline 2 & 2 \\ 3 & 1 \\ \hline 3 & 2 \\ 1 & 1 \end{array}\right], \quad \mathbf{X} = \left[\begin{array}{ccc} 1 & 0 & 1 \\ 0 & 1 & -1 \\ \hline 0 & 1 & 1 \\ -1 & -1 & -1 \\ \hline -1 & -1 & 1 \\ 1 & 0 & -1 \end{array}\right].$$

The three choice sets are separated by horizontal lines. Each row specifies an alternative, and the column dimension of the design matrix $\mathbf{X}$ corresponds to the number of parameter values k. Here, k = 3. We computed the Bayesian criterion values as we did in Equations 2, 3, 4, and 7. For the sake of illustration, we use only three prior parameters $\boldsymbol{\beta}^r = [\beta_{11}^r, \beta_{12}^r, \beta_2^r]'$, r = 1, 2, 3, randomly drawn from $\pi(\boldsymbol{\beta})$ =

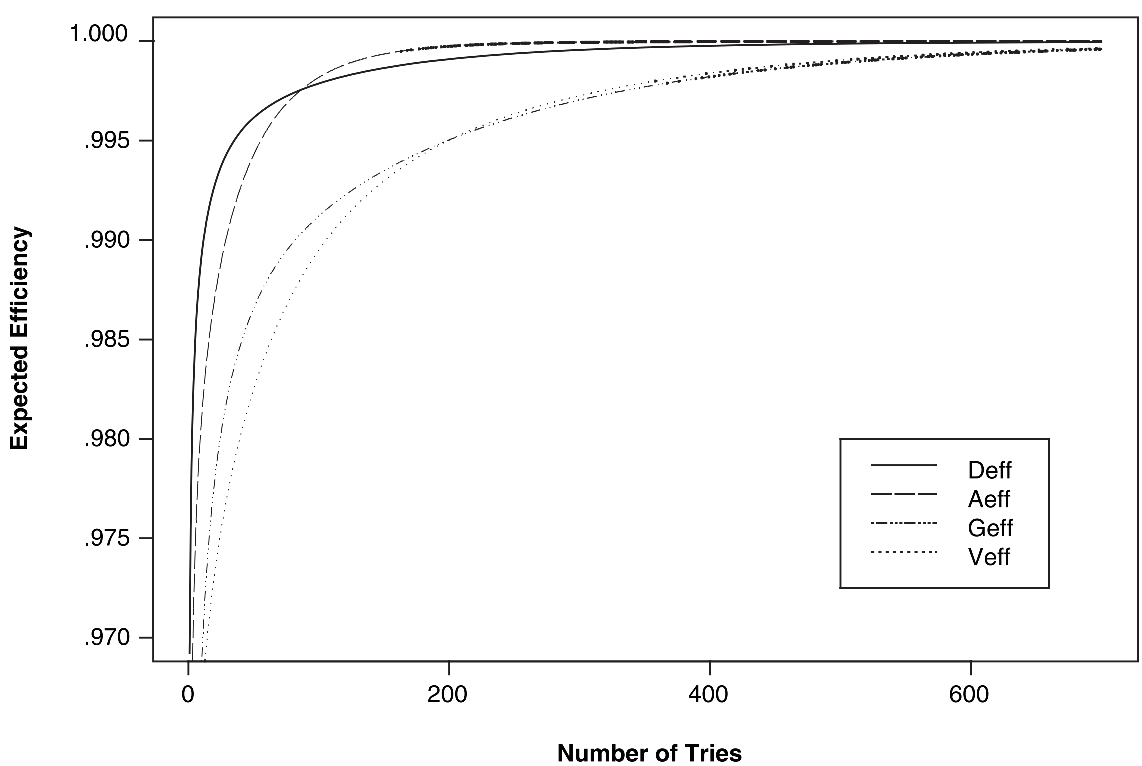
Figure 3

ESTIMATED EXPECTED EFFICIENCIES OF THE TWO- AND THREE-ALTERNATIVE D_B-, A_B-, G_B-, AND V_B-OPTIMAL DESIGNS

A: Two-Alternative Designs



B: Three-Alternative Designs



$N(\boldsymbol{\beta}|\boldsymbol{\beta}_0,\boldsymbol{\Sigma}_0)$, where $\boldsymbol{\beta}_0 = [-1, 0, -1]'$ and $\boldsymbol{\Sigma}_0$ is the 3×3 identity matrix. For each of these parameters, we compute the local D_P^r-, A_P^r-, G_P^r-, and V_P^r-criterion values and subsequently average them to obtain the Bayesian values.

We compute the information matrix $\mathbf{I}(\mathbf{X},\boldsymbol{\beta}^r)$ as we do in Equation 1 by taking N = 1 so that $\mathbf{I}(\mathbf{X},\boldsymbol{\beta}^r) = \Sigma_{s=1}^3 \mathbf{X}_s'(\mathbf{P}_s - \mathbf{p}_s\mathbf{p}_s')\mathbf{X}_s$ with choice sets given by

$$\text{(A2)} \quad \mathbf{X}_1 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}, \mathbf{X}_2 = \begin{bmatrix} 0 & 1 & 1 \\ -1 & -1 & -1 \end{bmatrix}, \mathbf{X}_3 = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix},$$

where $\mathbf{p}_s$ is the vector of probabilities in choice set s and $\mathbf{P}_s$ is the corresponding diagonal matrix. As a first draw, we have $\boldsymbol{\beta}^1 = [.805, -.080, -.603]'$. Using the multinomial logit model, we obtain the following for choice set s = 1:

$$\text{(A3)} \quad \mathbf{p}_1 = \begin{bmatrix} .420 \\ .580 \end{bmatrix}, \mathbf{P}_1 - \mathbf{p}_1\mathbf{p}_1' = .244 \times \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \text{ and}$$

$$\text{(A4)} \quad \mathbf{X}_1'(\mathbf{P}_1 - \mathbf{p}_1\mathbf{p}_1')\mathbf{X}_1 = \begin{bmatrix} .244 & -.244 & .487 \\ -.244 & .244 & -.487 \\ .487 & -.487 & .975 \end{bmatrix}.$$

Repeating the computations for choice sets 2 and 3 and summing the three matrices yields the information matrix pertaining to $\boldsymbol{\beta}^1$:

$$\text{(A5)} \quad \mathbf{I}(\mathbf{X},\boldsymbol{\beta}^1) = \begin{bmatrix} .703 & .333 & .721 \\ .333 & 1.226 & .324 \\ .721 & .324 & 2.128 \end{bmatrix}.$$

The local D_P^1-criterion value then becomes

$$\text{(A6)} \quad D_P^1 = \left\{\det\left(\mathbf{I}^{-1}(\mathbf{X},\boldsymbol{\beta}^1)\right)\right\}^{1/3} = .986,$$

and the local A_P^1-criterion value becomes

$$\text{(A7)} \quad A_P^1 = \mathrm{tr}\left(\mathbf{I}^{-1}(\mathbf{X},\boldsymbol{\beta}^1)\right) = 4.057.$$

To obtain the local G_P^1- and V_P^1-criterion values, we compute the prediction variances over the design region χ that consists of all possible choice sets of size two. For our small example, there are $3 \times 2 = 6$ candidate profiles, so that χ comprises $Q = \binom{6}{2} = 15$ choice sets, or 30 profiles. For each of these profiles, we compute the $\mathbf{c}$ vector according to Equation 6:

$$\text{(A8)} \quad \chi = \begin{bmatrix} \mathbf{x}_{11} \\ \mathbf{x}_{21} \\ \vdots \\ \mathbf{x}_{1,15} \\ \mathbf{x}_{2,15} \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & 1 \\ & \vdots & \\ -1 & -1 & -1 \\ -1 & -1 & 1 \end{bmatrix} \begin{matrix} \to & p_{11} = .770 & \to \\ \to & p_{21} = .230 & \to \\ & & \\ \to & p_{1,15} = .770 & \to \\ \to & p_{2,15} = .230 & \to \end{matrix}$$

$$\mathbf{c}(\mathbf{x}_{11}) = [0,\ 0,\ -.354]',$$
$$\mathbf{c}(\mathbf{x}_{21}) = [0,\ 0,\ .354]',$$

$$\mathbf{c}(\mathbf{x}_{1,15}) = [0,\ 0,\ -.354]',$$
$$\mathbf{c}(\mathbf{x}_{2,15}) = [0,\ 0,\ .354]'.$$

Then, we derive the local G_P^1- and V_P^1-criterion values as

$$\text{(A9)} \quad G_P^1 = \max\begin{pmatrix} \mathbf{c}'(\mathbf{x}_{11})\mathbf{I}^{-1}(\mathbf{X},\boldsymbol{\beta}^1)\mathbf{c}(\mathbf{x}_{11}) = .090 \\ \mathbf{c}'(\mathbf{x}_{21})\mathbf{I}^{-1}(\mathbf{X},\boldsymbol{\beta}^1)\mathbf{c}(\mathbf{x}_{21}) = .090 \\ \vdots \\ \mathbf{c}'(\mathbf{x}_{1,15})\mathbf{I}^{-1}(\mathbf{X},\boldsymbol{\beta}^1)\mathbf{c}(\mathbf{x}_{1,15}) = .090 \\ \mathbf{c}'(\mathbf{x}_{2,15})\mathbf{I}^{-1}(\mathbf{X},\boldsymbol{\beta}^1)\mathbf{c}(\mathbf{x}_{2,15}) = .090 \end{pmatrix} = .338, \text{ and}$$

$$\text{(A10)} \quad V_P^1 = \mathrm{avg}\begin{pmatrix} \mathbf{c}'(\mathbf{x}_{11})\mathbf{I}^{-1}(\mathbf{X},\boldsymbol{\beta}^1)\mathbf{c}(\mathbf{x}_{11}) = .090 \\ \mathbf{c}'(\mathbf{x}_{21})\mathbf{I}^{-1}(\mathbf{X},\boldsymbol{\beta}^1)\mathbf{c}(\mathbf{x}_{21}) = .090 \\ \vdots \\ \mathbf{c}'(\mathbf{x}_{1,15})\mathbf{I}^{-1}(\mathbf{X},\boldsymbol{\beta}^1)\mathbf{c}(\mathbf{x}_{1,15}) = .090 \\ \mathbf{c}'(\mathbf{x}_{2,15})\mathbf{I}^{-1}(\mathbf{X},\boldsymbol{\beta}^1)\mathbf{c}(\mathbf{x}_{2,15}) = .090 \end{pmatrix} = .170.$$

Similar computations for a second draw, $\boldsymbol{\beta}^2 = [-2.083, 2.238, -1.624]'$, yield

$$\text{(A11)} \quad \begin{aligned} D_P^2 &= 5.562, \\ A_P^2 &= 486.821, \\ G_P^2 &= 7.950, \\ V_P^2 &= 1.163, \end{aligned}$$

and for a third draw, $\boldsymbol{\beta}^3 = [-.486, -.087, -1.594]'$, we obtain

$$\text{(A12)} \quad \begin{aligned} D_P^3 &= 3.930, \\ A_P^3 &= 17.804, \\ G_P^3 &= 1.371, \\ V_P^3 &= .462. \end{aligned}$$

Finally, we average the local criterion values over the three draws to obtain the Bayesian values:

$$\text{(A13)} \quad \begin{aligned} D_B(\mathbf{X}) &= 3.493, \\ A_B(\mathbf{X}) &= 169.561, \\ G_B(\mathbf{X}) &= 3.219, \\ V_B(\mathbf{X}) &= .598. \end{aligned}$$

APPENDIX B

If T refers to the number of tries for the algorithm, the efficiency E_t of a design $\mathbf{X}_t$, t = 1, ..., T, generated by try t of the algorithm is given by

$$\text{(B1)} \quad E_t = \frac{B(\mathbf{X}^*)}{B(\mathbf{X}_t)},$$

where B represents the D_B-, A_B-, G_B-, or V_B-criterion value of a design and $\mathbf{X}^*$ is the optimal design according to that criterion.

Assume that for a large number of tries, T, we obtain G distinct designs $\mathbf{X}_1, ..., \mathbf{X}_G$, with efficiencies $E_1 > ... > E_G$ in terms of a particular optimality criterion. As such, $\mathbf{X}_1$ is the best design, and an estimate of the probability of finding $\mathbf{X}_1$ in T tries, for example, π_1, is given by the number of times $\mathbf{X}_1$ is found divided by T. Correspondingly, if $\pi_2, ..., \pi_G$

refer to the probabilities of finding $\mathbf{X}_2, ..., \mathbf{X}_G$ in T tries, the estimated expected efficiency from T tries is given by

$$\text{(B2)} \quad E(\text{efficiency}) = \sum_{i=1}^{G-1}\left\{\left(\sum_{j=i}^{G}\pi_j\right)^T - \left(\sum_{j=i+1}^{G}\pi_j\right)^T\right\}E_i + \pi_G^T E_G.$$

The mathematical derivation underlying this expression can be retrieved in the work of Trinca and Gilmour (2000), who introduce the estimated expected efficiency in the context of block designs.

REFERENCES

Anderson, Donald A. and James B. Wiley (1992), "Efficient Choice Set Designs for Estimating Availability Cross-Effects Models," *Marketing Letters*, 3 (4), 357–70.

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

Atkinson, Anthony C. and Alexander N. Donev (1992), *Optimum Experimental Designs*. Oxford: Clarendon Press.

——— and Linda M. Haines (1996), "Designs for Nonlinear and Generalized Linear Models," in *Handbook of Statistics: Design and Analysis of Experiments*, Vol. 13, Subir Ghosh and Calyampudi R. Rao, eds. Amsterdam: Elsevier, 437–75.

Bunch, David S., Jordan J. Louviere, and Donald A. Anderson (1996), "A Comparison of Experimental Design Strategies for Multinomial Logit Models: The Case of Generic Attributes," working paper, Graduate School of Management, University of California, Davis.

Burgess, Leonie and Deborah J. Street (2003), "Optimal Designs for 2^k Choice Experiments," *Communications in Statistics: Theory and Methods*, 32 (11), 2185-2206.

——— and ——— (2005), "Optimal Designs for Choice Experiments with Asymmetric Attributes," *Journal of Statistical Planning and Inference*, 134 (1), 288–301.

Carlsson, Fredrik and Peter Martinsson (2003), "Design Techniques for Stated Preference Methods in Health Economics," *Health Economics*, 12 (4), 281–94.

Carson, R.T., J.J. Louviere, D.A. Anderson, P. Arabie, D.S. Bunch, D.A. Hensher, R.M. Johnson, W.F. Kuhfeld, D. Steinberg, J. Swait, H. Timmermans, and J.B. Wiley (1994), "Experimental Analysis of Choice," *Marketing Letters*, 5 (4), 351–67.

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

Cook, R. Dennis and Christopher J. Nachtsheim (1980), "A Comparison of Algorithms for Constructing Exact D-Optimal Designs," *Technometrics*, 22 (3), 315–24.

Fedorov, Valeri V. (1972), *Theory of Optimal Experiments*. New York: Academic Press.

Goos, Peter (2002), *The Optimal Design of Blocked and Split-Plot Experiments*. New York: Springer.

Grossmann, Heiko, Heinz Holling, and Rainer Schwabe (2002), "Advances in Optimum Experimental Design for Conjoint Analysis and Discrete Choice Models," in *Advances in Econometrics, Econometric Models in Marketing*, Vol. 16, Philip H. Franses and Alan L. Montgomery, eds. Amsterdam: JAI Press, 93–117.

Huber, Joel and Klaus Zwerina (1996), "The Importance of Utility Balance in Efficient Choice Designs," *Journal of Marketing Research*, 33 (August), 307–317.

Kuhfeld, Warren F. and Randall D. Tobias (2005), "Large Factorial Designs for Product Engineering and Marketing Research Applications," *Technometrics*, 47 (2), 132–41.

———, ———, and Mark Garratt (1994), "Efficient Experimental Design with Marketing Research Applications," *Journal of Marketing Research*, 31 (November), 545–57. (Updated version [2005] available at http://support.sas.com/techsup/tnote/tnote_stat.html.)

Lazari, Andreas G. and Donald A. Anderson (1994), "Designs of Discrete Choice Experiments for Estimating Both Attribute and Availability Cross Effects," *Journal of Marketing Research*, 31 (August), 375–83.

Louviere, Jordan J., Deborah J. Street, and Leonie Burgess (2003), "A 20+ Years Retrospective on Choice Experiments," in *Marketing Research and Modeling: Progress and Prospects*, Ch. 8, Yoram Wind and Paul E. Green, eds. New York: Kluwer Academic Press, 201–214.

——— and George Woodworth (1983), "Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data," *Journal of Marketing Research*, 20 (November), 350–67.

McFadden, Daniel (1974), "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics*, Paul Zarembka, ed. New York: Academic Press, 105–142.

Meyer, Ruth K. and Christopher J. Nachtsheim (1995), "The Coordinate Exchange Algorithm for Constructing Exact Optimal Experimental Designs," *Technometrics*, 37 (1), 60–69.

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

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

Silvey, Samuel D. (1980), *Optimal Design*. London: Chapman and Hall.

Street, Deborah J., David S. Bunch, and Beverly J. Moore (2001), "Optimal Designs for 2^k Paired Comparison Experiments," *Communications in Statistics: Theory and Methods*, 30 (10), 2149–71.

——— and Leonie Burgess (2004), "Optimal and Near Optimal Pairs for the Estimation of Effects in 2-Level Choice Experiments," *Journal of Statistical Planning and Inference*, 118 (1–2), 185–99.

Swait, Joffre and Wiktor Adamowicz (2001), "The Influence of Task Complexity on Consumer Choice: A Latent Class Model of Decision Strategy Switching," *Journal of Consumer Research*, 28 (June), 135–48.

Trinca, Luzia A. and Steven G. Gilmour (2000), "An Algorithm for Arranging Response Surface Designs in Small Blocks," *Computational Statistics and Data Analysis*, 33 (1), 25–43.

Zwerina, Klaus, Joel Huber, and Warren F. Kuhfeld (1996), "A General Method for Constructing Efficient Choice Designs," working paper, Fuqua School of Business, Duke University. (Updated version [2005] available at http://support.sas.com/techsup/tnote/tnote_stat.html.)