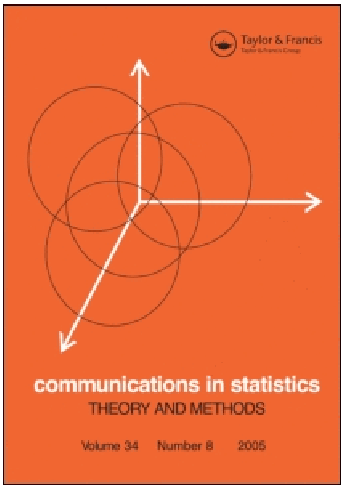

Communications in Statistics - Theory and Methods
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597238

Optimal Designs for 2<i>k</i> Choice Experiments

Leonie Burgess [a]; Deborah J. Street [a]
[a] Department of Mathematical Sciences, University of Technology Sydney, NSW, Australia

Online publication date: 27 August 2003

To cite this Article Burgess, Leonie and Street, Deborah J.(2003) 'Optimal Designs for 2<i>k</i> Choice Experiments', Communications in Statistics - Theory and Methods, 32: 11, 2185 — 2206
To link to this Article: DOI: 10.1081/STA-120024475
URL: http://dx.doi.org/10.1081/STA-120024475

MARCEL DEKKER, INC. • 270 MADISON AVENUE • NEW YORK, NY 10016

COMMUNICATIONS IN STATISTICS
Theory and Methods
Vol. 32, No. 11, pp. 2185–2206, 2003

Optimal Designs for 2^k Choice Experiments

Leonie Burgess* and Deborah J. Street

Department of Mathematical Sciences, University of Technology Sydney, NSW, Australia

ABSTRACT

In this article we establish the choice sets in the D-optimal design for a choice experiment for testing main effects and for testing main effects and two-factor interactions, when there are k attributes, each with two levels, for choice set size m. We also give a method to construct optimal and near-optimal designs with small numbers of choice sets.

Key Words: Paired comparisons; Multiple comparisons; Bradley–Terry model; Multinomial logit model; Fractional factorial designs; Orthogonal main effect plans.

*Correspondence: Leonie Burgess, Department of Mathematical Sciences, University of Technology Sydney, Broadway, NSW 2007, Australia; E-mail: leonie.burgess@uts.edu.au.

DOI: 10.1081/STA-120024475 0361-0926 (Print); 1532-415X (Online)
 www.dekker.com

1. INTRODUCTION

Stated choice experiments have become an established method of obtaining information about choices in various areas including marketing, transport, environmental resource economics, and public welfare analysis; see Louviere et al. (2000) for an overview of the area.

We begin with a simple example. Suppose that we are interested in the effect of three attributes (factors) on the choice of car loans. The attributes are *provider* with levels *bank* and *credit union*, *early repayment penalty* with levels \$250 and \$400 and *establishment fee* with levels \$190 and \$100. Using these attributes we can describe eight different car loans (profiles or treatment combinations). One such is provider = bank, early repayment penalty = \$250, and establishment fee = \$190 for example. We now use these profiles to construct some choice sets. The first choice set might be (bank, \$400, \$100), (bank, \$250, \$190), and (credit union, \$250, \$100) and we would ask each respondent to say which of these three car loans they prefer. A choice experiment consists of a set of choice sets. Each respondent is shown all of the choice sets in the experiment, one after the other, and for each choice set they are asked to choose one of the options presented. Provided that the choice experiment has been correctly designed, these responses can be used to estimate the effect of each of the attributes on car loan choice and to estimate the effect of the interaction of any two of the attributes on the choice of car loan.

In general we will describe the products to be compared by k attributes and will show the respondents choice sets of size m. We will insist that respondents choose one of the options in each choice set (termed *forced choice* in the literature). This is the appropriate setting when you must make a choice, for instance when you must buy a car and do not have enough cash to do so and hence must borrow money, or when faced with making a choice about how to commute to work or where to stay while on an overnight business trip.

The aim of this article is to give optimal choice experiments when there are k binary attributes describing each possible choice and choice sets are of size m. The goal of the experiment may be to estimate main effects or to estimate main effects and two-factor interactions.

2. PRELIMINARY RESULTS

Several results on optimal choice pairs for forced choice experiments involving two level attributes have appeared in the literature.

In all cases that have appeared to date, forced choice has been assumed and all pairs with i attributes different have been assumed to appear equally often. Results have been given by El-Helbawy and Ahmed (1984), El-Helbawy et al. (1994), Street et al. (2001), and van Berkum (1987) on the optimal pairs from a complete factorial. As a result of the work in these articles we know that the optimal pairs for the estimation of main effects are the pairs with all attributes different and the optimal pairs for the estimation of main effects and two-factor interactions are pairs with about half the attributes different.

As choice experiments based on a complete factorial rapidly become very large, Street and Burgess (to appear) looked at the use of fractional factorial designs for the construction of optimal forced-choice paired comparison experiments. They established that fold-over pairs constructed from a resolution 3 fractional factorial design were D-optimal. For the estimation of main effects and two factor interactions they gave sets of generators which could be added, in turn, to the profiles in a resolution 5 fractional factorial design to give a set of pairs that were optimal or near-optimal.

In this article we extend the work on pairs to forced-choice experiments with larger choice set sizes, although we still retain the restriction that all attributes have only two levels. The only work that we are aware of that considers the optimal design of choice sets with more than two alternatives is given in Bunch et al. (1996). Various strategies for constructing choice sets of size two or more, for attributes with two or more levels, were described and compared in the article. The most efficient of the strategies considered starts with a fractional factorial design, then shifts each attribute level to a different level to obtain the next profile in the choice set. This shifting process is repeated until the desired choice set size is reached, as long as the choice set size is no greater than the number of attribute levels. Hence for 2-level attributes this method can only construct pairs.

Here we use a multinomial logit (MNL) model and we generalize the results in Bradley (1955), Bradley and Terry (1952), and Pendergrass and Bradley (1960) to obtain the information matrix for this model. We then extend the techniques in Street et al. (2001) to give the form of the D-optimal designs for any choice set size, when estimating main effects only, and when estimating main effects and two-factor interactions. We also extend the constructions in Street and Burgess (to appear) to obtain optimal or near-optimal designs with small numbers of choice sets, for any choice set size for these two situations.

3. DIFFERENCE VECTORS

In choice experiments, we are interested in the number of attributes with equal levels and the number with different levels in the choice set, as this indicates how efficiently main effects and interaction effects can be estimated (Street et al., 2001). We will denote the levels of the attributes by 0 and 1 in general. In a choice set of size m there are $\binom{m}{2}$ pairs of profiles (or treatments) in the choice set and we record the number of attributes different for each pair in the choice set in a *difference vector* $\mathbf{v} = (d_1, d_2, \ldots, d_{m(m-1)/2})$, where $1 \leq d_i \leq k$; that is, no repeated treatments are allowed. We define d_1 to be the number of attributes different (or difference) between the first and second treatments in the choice set; d_2 is the difference between the first and third treatments, and so on.

Since we can write the treatments in the choice set in any order, the order of the d_i in $\mathbf{v}$ is not important, so we assume that any difference vector has $d_1 \leq d_2 \leq \cdots \leq d_{m(m-1)/2}$. For example, for $k = 3$ and $m = 3$ the choice sets $(000, 001, 110)$ and $(011, 100, 101)$ have difference vectors $(1, 2, 3)$ and $(3, 2, 1)$ respectively. These difference vectors are considered to be the same and are written as $(1, 2, 3)$.

We now establish the upper bound for the sum of the differences in a difference vector.

Lemma 1. *For a particular difference vector* $\mathbf{v}$, *for a given* m *and* $k \geq \ell$, *where* $2^{\ell-1} < m \leq 2^{\ell}$ *the least upper bound for the sum of the differences is* $(m^2 - 1)k/4$ *for* m *odd, and* $m^2k/4$ *for* m *even.*

Proof. Write the treatments in the choice set as the rows of an $m \times k$ array. Then for each column of length m the maximum contribution to $\sum_{i=1}^{m(m-1)/2} d_i$ comes by having half the entries 1 and half 0 if m is even, or $(m-1)/2$ entries 1 and $(m+1)/2$ entries 0 (or the other way round) if m is odd. To get m distinct rows we must have at least ℓ columns where $2^{\ell-1} < m \leq 2^{\ell}$. Thus we get m distinct rows of ℓ columns and with a maximum difference by writing down the rows in foldover pairs. (One treatment combination is said to be the foldover of another if the second one is obtained from the first by changing all 0s to 1s and all 1s to 0s.) Thus we write

$$\begin{array}{ccccc} 0 & 0 & \ldots & 0 & 0 \\ 1 & 1 & \ldots & 1 & 1 \\ 0 & 0 & \ldots & 0 & 1 \\ 1 & 1 & \ldots & 1 & 0 \end{array}$$

and so on. If m is odd we have $(m-1)/2$ foldover pairs and one extra row, which can be any treatment not already used. It does not matter which particular treatments are used to construct the rows. Because the rows appear in foldover pairs, half the entries are 1 and half are 0 (m even) and so $\sum_{i=1}^{m(m-1)/2} d_i$ is a maximum. Since these ℓ columns guarantee that the rows are distinct, larger values of k can be obtained by writing down any columns of maximum difference. The result follows.

For example, for $k=3$ and $m=3$ the possible difference vectors are $(1,1,2)$, $(1,2,3)$, and $(2,2,2)$, with sums 4, 6, and 6 respectively. The upper bound is $(3^2-1)3/4=6$. The value of ℓ is 2 and a set of rows constructed as in the lemma is

$$\begin{matrix} 0 & 0 \\ 1 & 1 \\ 0 & 1 \end{matrix}$$

To get a choice set which meets the bound we can now adjoin any column with a 1 and two 0s. So we might get $(000, 110, 011)$ which has difference vector $(2,2,2)$.

For particular values of m and k there can be several difference vectors; these are denoted by $\mathbf{v}_j$. We now define three scalars which are needed subsequently. We define $c_{\mathbf{v}_j}$ to be the number of choice sets containing the treatment $00\ldots0$ with the difference vector $\mathbf{v}_j$, and define $x_{\mathbf{v}_j;i}$ to be the number of times the difference i appears in the difference vector $\mathbf{v}_j$. We define $a_{\mathbf{v}_j}$ to be an indicator variable, where $a_{\mathbf{v}_j}=0$ if no choice sets have the difference vector $\mathbf{v}_j$ and $a_{\mathbf{v}_j}=1/(\text{the total number of choice sets in the experiment})$ if there are choice sets with the difference vector $\mathbf{v}_j$. At least one of the $a_{\mathbf{v}_j}$ values must be nonzero otherwise the experiment contains no choice sets. This is similar to the $a_{k,i}$ in Street et al. (2001), with the added restriction that a particular choice set may only appear once, or not at all.

Example 1. For $m=3$, $k=3$ and using all treatments from the complete 2^3 factorial, there are $\binom{2^3}{3}=56$ distinct choice sets of size 3, 24 with difference vector $(1,1,2)$, 24 with difference vector $(1,2,3)$ and 8 with difference vector $(2,2,2)$. Let $\mathbf{v}_1=(1,1,2)$, $\mathbf{v}_2=(1,2,3)$, and $\mathbf{v}_3=(2,2,2)$. Now consider just the triples containing the treatment 000. Thus we have $c_{\mathbf{v}_1}=c_{\mathbf{v}_2}=9$ and $c_{\mathbf{v}_3}=3$, since, for instance, $(000, 011, 101), (000, 011, 110)$, and $(000, 101, 110)$ are the three choice sets with $m=3$ which contain 000 and which have difference vector $(2,2,2)$.

4. THE MODEL AND THE INFORMATION MATRIX

Consider an experiment in which there are N choice sets of m treatments, of which $n_{i_1,i_2,\ldots,i_m}$ compare the specific treatments $T_{i_1}, T_{i_2}, \ldots, T_{i_m}$, where $n_{i_1,i_2,\ldots,i_m} = 1$ if $(T_{i_1}, T_{i_2}, \ldots, T_{i_m})$ is a choice set and is 0 otherwise. Then

$$N = \sum_{i_1<i_2<\cdots<i_m} n_{i_1,i_2,\ldots,i_m}.$$

In choice experiments we define the parameters

$$\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_{2^k})$$

associated with 2^k treatments $T_1, T_2, \ldots, T_{2^k}$ and in this article we consider the multinomial logit (MNL) model (see, for example, Louviere et al. (2000)). In the MNL model given a choice set which contains m treatments $T_{i_1}, T_{i_2}, \ldots, T_{i_m}$ the probability that treatment T_{i_1} is preferred to the other $m-1$ treatments in the choice set is

$$P(T_{i_1} > T_{i_2}, \ldots, T_{i_m}) = \frac{\pi_{i_1}}{\sum_{j=1}^{m} \pi_{i_j}}$$

for $i_j = 1, 2, \ldots, 2^k$ and no two treatments are the same. Choices made in one choice set do not affect choices made in any other choice set. If $m = 2$ this is just the Bradley–Terry model (see Bradley and Terry (1952)). This is the model considered in Street et al. (2001).

We let $\Lambda_k(\boldsymbol{\pi})$ be the matrix of second derivatives of the likelihood function, where $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_{2^k})$. For choice sets of size 2, Street et al. (2001) have provided a general form of $\Lambda_k(\pi)$ and information matrix $C_k = B_k \Lambda_k B_k'/2^k$, where B_k is the $(2^k - 1) \times 2^k$ matrix of contrasts associated with a 2^k factorial design. We now extend this work to obtain the general form of $\Lambda_k(\boldsymbol{\pi})$, for any choice set size, and hence the general form of C_k. We use the method in Bradley (1955) and Pendergrass and Bradley (1960) to obtain the form of the entries in $\Lambda_k(\boldsymbol{\pi})$.

Let $w_{i_1,i_2,\ldots,i_m}$ be an indicator variable where

$$w_{i_1,i_2,\ldots,i_m} = \begin{cases} 1 & \text{if } T_{i_1} > T_{i_2}, T_{i_3}, \ldots, T_{i_m}, \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$E(w_{i_1,i_2,\ldots,i_m}) = \frac{\pi_{i_1}}{\sum_{j=1}^{m} \pi_{i_j}} \quad \text{and} \quad \mathrm{Var}(w_{i_1,i_2,\ldots,i_m}) = \frac{\pi_{i_1} \sum_{j=2}^{m} \pi_{i_j}}{(\sum_{j=1}^{m} \pi_{i_j})^2}.$$

We let w_{i_1} be the number of times treatment T_{i_1} is preferred to treatments $T_{i_2}, \ldots, T_{i_m}$ considered over all choice sets in which T_{i_1} appears. Then

$$w_{i_1} = \sum_{i_2<i_3<\cdots<i_m}{}' w_{i_1, i_2, \ldots, i_m}$$

where the summation is over $i_2 < i_3 < \cdots < i_m$ and $i_j \neq i_1$ for $j = 2, \ldots, m$.

It follows that

$$E(w_{i_1}) = \pi_{i_1} \sum_{i_2<i_3<\cdots<i_m}{}' \frac{1}{\sum_{j=1}^{m} \pi_{i_j}},$$

$$\text{Var}(w_{i_1}) = \pi_{i_1} \sum_{i_2<i_3<\cdots<i_m}{}' \frac{\sum_{j=2}^{m} \pi_{i_j}}{(\sum_{j=1}^{m} \pi_{i_j})^2}$$

and

$$\text{Cov}(w_{i_1}, w_{i_2}) = \frac{-\pi_{i_1}\pi_{i_2}}{(\sum_{j=1}^{m} \pi_{i_j})^2}.$$

We define $\lambda_{i_1, i_2, \ldots, i_m} = n_{i_1, i_2, \ldots, i_m}/N$ and the entries of Λ_k are given by

$$\Lambda_{i_1, i_1} = \pi_{i_1} \sum_{i_2<i_3<\cdots<i_m}{}' \frac{\lambda_{i_1, i_2, \ldots, i_m} \sum_{j=2}^{m} \pi_{i_j}}{(\sum_{j=1}^{m} \pi_{i_j})^2}$$

and

$$\Lambda_{i_1, i_2} = -\pi_{i_1}\pi_{i_2} \sum_{i_3<i_4<\cdots<i_m}{}' \frac{\lambda_{i_1, i_2, \ldots, i_m}}{(\sum_{j=1}^{m} \pi_{i_j})^2}.$$

If we assume that $\pi_{i_1} = \pi_{i_2} = \cdots = \pi_{i_{2^k}} = \pi_0$, say, (that is, all treatments are equally attractive, the usual null hypothesis) then

$$\Lambda_{i_1, i_1} = \frac{m-1}{m^2} \sum_{i_2<i_3<\cdots<i_m}{}' \lambda_{i_1, i_2, \ldots, i_m}$$

and

$$\Lambda_{i_1, i_2} = -\frac{1}{m^2} \sum_{i_3<i_4<\cdots<i_m}{}' \lambda_{i_1, i_2, \ldots, i_m}.$$

The off-diagonal entry Λ_{i_1,i_2} is by definition the proportion of choice sets in which treatments T_{i_1} and T_{i_2} appear together, multiplied by $-1/m^2$. Since by definition the rows, and columns, of Λ_k must sum to zero, the diagonal entry in a row (or column) must equal the negative sum of the off-diagonal entries in that row (or column).

In Street et al. (2001) the general form of the Λ_k matrix for $m = 2$ was given as a linear combination of the identity matrix of order 2^k and some $D_{k,i}$ matrices, where $D_{k,i}$ is a $(0, 1)$ matrix of order 2^k with a 1 in position (x, y) if treatment combinations x and y have i attributes with different levels. Using this definition and the same method as Street et al. (2001), we find that the general form of the Λ_k matrix for any choice set size is given by

$$\Lambda_k = \frac{m-1}{m^2} z I_{2^k} - \frac{1}{m^2} \sum_{i=1}^{k} y_i D_{k,i}$$

where

$$y_i = \frac{2}{m} \binom{k}{i}^{-1} \sum_j c_{\mathbf{v}_j} x_{\mathbf{v}_j;i} a_{\mathbf{v}_j}$$

and

$$z = \sum_j c_{\mathbf{v}_j} a_{\mathbf{v}_j} = \frac{1}{m-1} \sum_{i=1}^{k} \binom{k}{i} y_i.$$

The summations over j are over all possible difference vectors $\mathbf{v}_j$ for that particular value of k.

The y_i values represent the linear combination of the $c_{\mathbf{v}_j} x_{\mathbf{v}_j;i} a_{\mathbf{v}_j}$ values for those difference vectors $\mathbf{v}_j$ in which the difference i appears.

Example 2. Consider $m = 3$, $k = 3$ from Example 1. Using all treatments from the complete 2^3 factorial there are three possible difference vectors $\mathbf{v}_1 = (1, 1, 2)$, $\mathbf{v}_2 = (1, 2, 3)$, and $\mathbf{v}_3 = (2, 2, 2)$. Since the difference 1 appears in $\mathbf{v}_1$ twice and in $\mathbf{v}_2$ once, y_1 will contain $a_{\mathbf{v}_1}$ twice and $a_{\mathbf{v}_2}$ once. Similarly, for y_2 the difference 2 appears in all three difference vectors, once in $\mathbf{v}_1$ and $\mathbf{v}_2$, and three times in $\mathbf{v}_3$. Finally, the difference 3 appears in $\mathbf{v}_2$ only. Thus we have

$$\begin{aligned}
y_1 &= 2(2c_{\mathbf{v}_1} a_{\mathbf{v}_1} + c_{\mathbf{v}_2} a_{\mathbf{v}_2})/9 = 4a_{\mathbf{v}_1} + 2a_{\mathbf{v}_2}, \\
y_2 &= 2(c_{\mathbf{v}_1} a_{\mathbf{v}_1} + c_{\mathbf{v}_2} a_{\mathbf{v}_2} + 3c_{\mathbf{v}_3} a_{\mathbf{v}_3})/9 = 2a_{\mathbf{v}_1} + 2a_{\mathbf{v}_2} + 2a_{\mathbf{v}_3}, \\
y_3 &= 2(c_{\mathbf{v}_2} a_{\mathbf{v}_2})/3 = 6a_{\mathbf{v}_2}.
\end{aligned}$$

5. DESIGNS FOR MAIN EFFECTS ONLY

We now evaluate the $k \times k$ principal minor of $C_k = B_k \Lambda_k B_k'/2^k$, associated with the main effects and determine the form of the *D*-optimal designs. We let $B_{k,M}$ be the rows of B_k that correspond to main effects. In Street et al. (2001) it is shown that

$$B_{k,M} D_{k,i} = \left[\binom{k-1}{i} - \binom{k-1}{i-1}\right] B_{k,M}$$

for all allowable *i*. The $k \times k$ principal minor of C_k, after normalizing the contrast matrix, is

$$\begin{aligned} C_{k,M} &= \frac{1}{2^k} B_{k,M} \Lambda_k(\pi) B_{k,M}' \\ &= \frac{2}{m^2} \sum_{i=1}^{k} y_i \binom{k-1}{i-1} I_k. \end{aligned}$$

Thus the determinant of $C_{k,M}$ is

$$\det(C_{k,M}) = \left[\frac{2}{m^2} \sum_{i=1}^{k} y_i \binom{k-1}{i-1}\right]^k.$$

To find the *D*-optimal design, we must maximize $\det(C_{k,M})$ subject to the constraint $2^k z/m = 1$.

Substituting for y_i gives

$$\det(C_{k,M}) = \left[\frac{4}{m^3 k} \sum_{i=1}^{k} i \left(\sum_j c_{\mathbf{v}_j} x_{\mathbf{v}_j;i} a_{\mathbf{v}_j}\right)\right]^k.$$

Recall that $a_{\mathbf{v}_j}$ is an indicator variable and at least one of these values must be nonzero. The following theorem establishes that the *D*-optimal design, for estimating main effects only, is one which consists of choice sets in which the sum of the differences attains the maximum value given in Lemma 1.

Theorem 1. *The D-optimal design for testing main effects only, when all other effects are assumed to be zero, is given by choice sets in which, for each* $\mathbf{v}_j$ *present,*

$$\sum_{i=1}^{m(m-1)/2} d_{ij} = \begin{cases} (m^2-1)k/4, & m \text{ odd}, \\ m^2 k/4, & m \text{ even}, \end{cases}$$

and there is at least one $\mathbf{v}_j$ *with a nonzero* $a_{\mathbf{v}_j}$; *that is, the choice set is nonempty.*

Proof. Since $x_{\mathbf{v}_j;i}$ denotes the number of times the difference i appears in the difference vector $\mathbf{v}_j$, multiplying by i and summing these is equivalent to summing the entries in $\mathbf{v}_j$. Hence

$$\sum_{i=1}^{k} i x_{\mathbf{v}_j;i} = \sum_{i=1}^{m(m-1)/2} d_{ij}.$$

Recall that $a_{\mathbf{v}_j} = (2^k \sum_j c_{\mathbf{v}_j}/m)^{-1}$. Then, substituting,

$$\det(C_{k,M}) = \left[\frac{\sum_j \left(\sum_{i=1}^{m(m-1)/2} d_{ij}\right) c_{\mathbf{v}_j}}{m^2 k 2^{k-2} \sum_j c_{\mathbf{v}_j}}\right]^k.$$

When m is even, it follows from Lemma 1 that $\sum_{i=1}^{m(m-1)/2} d_{ij} = m^2 k/4 - \ell_j$ for $\ell_j \geq 0$. Thus

$$\det(C_{k,M}) = \left[\frac{1}{2^k} - \frac{\sum_j \ell_j c_{\mathbf{v}_j}}{m^2 k 2^{k-2} \sum_j c_{\mathbf{v}_j}}\right]^k.$$

Since the $c_{\mathbf{v}_j}$ values are all positive, for m even, $\det(C_{k,M})$ is a maximum of $(1/2^k)^k$ when $\ell_j = 0$ for all j. Thus we obtain the maximum $\det(C_{k,M})$ when

$$\sum_{i=1}^{m(m-1)/2} d_{ij} = m^2 k/4.$$

Similarly when m is odd, from Lemma 1 $\sum_{i=1}^{m(m-1)/2} d_{ij} = (m^2-1)k/4 - \ell_j$ for $\ell_j \geq 0$. Then

$$\det(C_{k,M}) = \left[\frac{m^2-1}{m^2 2^k} - \frac{\sum_j \ell_j c_{\mathbf{v}_j}}{m^2 k 2^{k-2} \sum_j c_{\mathbf{v}_j}}\right]^k.$$

Since the $c_{\mathbf{v}_j}$ values are all positive, $\det(C_{k,M})$ is a maximum of $((m^2-1)/(m^2 2^k))^k$ when $\ell_j = 0$ for all j. For m odd, we obtain the maximum $\det(C_{k,M})$ when

$$\sum_{i=1}^{m(m-1)/2} d_{ij} = (m^2-1)k/4.$$

Example 3. Recall that for $m = 3$ and $k = 3$, there are three difference vectors $\mathbf{v}_1 = (1, 1, 2)$, $\mathbf{v}_2 = (1, 2, 3)$, and $\mathbf{v}_3 = (2, 2, 2)$. Theorem 1 states that the D-optimal designs have choice sets in which the entries in the difference vectors sum to $(m^2 - 1)k/4 = 2k = 6$. The difference vectors $\mathbf{v}_2$ and $\mathbf{v}_3$ have entries which sum to $2k$, so there are three D-optimal designs:

1. All 24 triples with difference vector $\mathbf{v}_2$ where $a_{\mathbf{v}_2} = 1/24$ and $a_{\mathbf{v}_1} = a_{\mathbf{v}_3} = 0$.
2. All 8 triples with difference vector $\mathbf{v}_3$ where $a_{\mathbf{v}_3} = 1/8$ and $a_{\mathbf{v}_1} = a_{\mathbf{v}_2} = 0$.
3. All 32 triples with difference vectors $\mathbf{v}_2$ and $\mathbf{v}_3$ where $a_{\mathbf{v}_2} = a_{\mathbf{v}_3} = 1/32$ and $a_{\mathbf{v}_1} = 0$.

The smallest of these D-optimal designs is the second one, consisting of the following eight triples, each with difference vector $\mathbf{v}_3$: (000,011,101), (000, 011, 110), (000, 101, 110), (001, 010, 100), (001, 010, 111), (001, 100, 111), (010, 100, 111), and (011, 101, 110).

6. DESIGNS FOR MAIN EFFECTS AND TWO-FACTOR INTERACTIONS

To find the D-optimal designs for estimating main effects and two-factor interactions for any choice set size m, we generalize the results of Street et al. (2001). We know the general form of C_k and now evaluate the $[k + \binom{k}{2}] \times [k + \binom{k}{2}]$ principal minor of C_k associated with the main effects and two-factor interactions. We let $B_{k,M}$ be the rows of B_k that correspond to main effects and we let $B_{k,F}$ be the rows of B_k that correspond to the two-factor interactions. The matrix associated with main effects and two-factor interactions is denoted by $B_{k,MF}$ and is the concatenation of $B_{k,M}$ and $B_{k,F}$.

In Street et al. (2001) it is established that

$$B_{k,F} D_{k,i} = \left[\binom{k-2}{i} - 2\binom{k-2}{i-1} + \binom{k-2}{i-2} \right] B_{k,F}$$

for all allowable i. Thus we can show that the $[k + \binom{k}{2}] \times [k + \binom{k}{2}]$ principal minor is

$$\begin{aligned} C_{k,MF} &= B_{k,MF} \Lambda(\pi) B'_{k,MF} \\ &= \begin{bmatrix} \frac{2}{m^2} \sum_{i=1}^{k} y_i \binom{k-1}{i-1} I_k & 0 \\ 0 & B_{k,F} \Lambda B'_{k,F} \end{bmatrix}. \end{aligned}$$

MARCEL DEKKER, INC. • 270 MADISON AVENUE • NEW YORK, NY 10016

Now, after normalizing the contrast matrix $B_{k,F}$,

$$B_{k,F}\Lambda = \frac{1}{\sqrt{2^k}} B_{k,F}\left[\frac{m-1}{m^2} z I_{2^k} - \frac{1}{m^2}\sum_{i=1}^{k} y_i D_{k,i}\right]$$
$$= \frac{4}{m^2\sqrt{2^k}}\sum_{i=1}^{k} y_i \binom{k-2}{i-1} B_{k,F}.$$

Thus

$$B_{k,F}\Lambda B'_{k,F} = \frac{4}{m^2}\sum_{i=1}^{k} y_i \binom{k-2}{i-1} I_{k(k-1)/2}.$$

So the determinant of $C_{k,MF}$ is

$$\det(C_{k,MF}) = \left[\frac{2}{m^2}\sum_{i=1}^{k} y_i\binom{k-1}{i-1}\right]^k \times \left[\frac{4}{m^2}\sum_{i=1}^{k} y_i\binom{k-2}{i-1}\right]^{k(k-1)/2}$$
$$= \left[\frac{4}{m^3 k}\sum_{i=1}^{k} i\left(\sum_j c_{\mathbf{v}_j} x_{\mathbf{v}_j;i} a_{\mathbf{v}_j}\right)\right]^k$$
$$\times \left[\frac{8\sum_{i=1}^{k} i(k-i)}{m^3 k(k-1)}\left(\sum_j c_{\mathbf{v}_j} x_{\mathbf{v}_j;i} a_{\mathbf{v}_j}\right)\right]^{k(k-1)/2}.$$

To find the D-optimal design, we need to maximize $\det(C_{k,MF})$ subject to the constraint that $2^k z/m = 1$.

Theorem 2. *The D-optimal design for testing main effects and two-factor interactions when all other effects are assumed to be zero, is given by*

$$y_i = \begin{cases} \frac{m(m-1)}{2^k}\binom{k+1}{k/2}^{-1} & k \text{ even and } i = k/2,\ k/2+1, \\ \frac{m(m-1)}{2^k}\binom{k}{(k+1)/2}^{-1} & k \text{ odd and } i = (k+1)/2, \\ 0 & \text{otherwise,} \end{cases}$$

when this results in nonzero y_i's that correspond to difference vectors that actually exist.

Proof. In Lemma 1 of Street et al. (2001) it was proved that for $m = 2$ the D-optimal design for testing main effects and two-factor interactions is given by

$$x_i = \begin{cases} \binom{k+1}{k/2}^{-1} & k \text{ even and } i = k/2,\ k/2+1, \\ \binom{k}{(k+1)/2}^{-1} & k \text{ odd and } i = (k+1)/2, \\ 0 & \text{otherwise.} \end{cases}$$

where $a_{k,i} = x_i/2^{k-1}$. In their proof of this lemma, the function $f = AB^{(k-1)/2}$, and therefore $\det(C_{k,MF})$, is maximized subject to the constraint $\sum_{i=1}^{k}\binom{k}{i}x_i = 1$. In Street et al. (2001)

$$A = 2^{k-1}\sum_{i=1}^{k}\binom{k-1}{i-1}a_{k,i} \quad \text{and} \quad B = 2^{k-1}\sum_{i=1}^{k}\binom{k-2}{i-1}a_{k,i}.$$

For choice sets of size m we let $x_i = 2^k y_i/(m(m-1))$ (suggested by Moore (2001) when $m = 3$) and we have

$$A = \frac{2^k}{m(m-1)}\sum_{i=1}^{k}\binom{k-1}{i-1}y_i \quad \text{and} \quad B = \frac{2^k}{m(m-1)}\sum_{i=1}^{k}\binom{k-2}{i-1}y_i.$$

The function $f = AB^{(k-1)/2}$, and therefore

$$\det(C_{k,MF}) = \left[\frac{2}{m^2}\sum_{i=1}^{k} y_i\binom{k-1}{i-1}\right]^k \times \left[\frac{4}{m^2}\sum_{i=1}^{k} y_i\binom{k-2}{i-1}\right]^{k(k-1)/2}$$

is maximized subject to the constraint $\sum_{i=1}^{k}\binom{k}{i}x_i = 1$ for the same x_i values given above for the $m = 2$ case. Then using $x_i = 2^k y_i/(m(m-1))$ we obtain the optimal designs in terms of the y_i values as required. The maximum value of the determinant at these y_i values is

$$\det(C_{k,MF}) = \begin{cases} \left(\dfrac{(m-1)(k+2)}{m(k+1)2^k}\right)^{k+k(k-1)/2} & k \text{ even}, \\ \left(\dfrac{(m-1)(k+1)}{mk2^k}\right)^{k+k(k-1)/2} & k \text{ odd}. \end{cases}$$

Example 4. For $m = 3$ and $k = 4$ the possible difference vectors are $\mathbf{v}_1 = (1, 1, 2)$, $\mathbf{v}_2 = (1, 2, 3)$, $\mathbf{v}_3 = (1, 3, 4)$, $\mathbf{v}_4 = (2, 2, 2)$, $\mathbf{v}_5 = (2, 2, 4)$, and $\mathbf{v}_6 = (2, 3, 3)$. Using Theorem 2, the D-optimal design is given by

$$y_i = \begin{cases} \dfrac{3 \times 2}{2^4}\binom{5}{2}^{-1} = \dfrac{6}{160} & i = 2, 3, \\ 0 & \text{otherwise}. \end{cases}$$

Now $\mathbf{v}_4$ and $\mathbf{v}_6$ are the only difference vectors containing all 2's, all 3's or a combination of 2's and 3's, so y_2 and y_3 are the only y_i values that are nonzero. We have

$$y_1 = 6a_{\mathbf{v}_1} + 6a_{\mathbf{v}_2} + 2a_{\mathbf{v}_3} = 0$$
$$y_2 = 2a_{\mathbf{v}_1} + 4a_{\mathbf{v}_2} + 4a_{\mathbf{v}_4} + 2a_{\mathbf{v}_5} + 2a_{\mathbf{v}_6} = 6/160$$
$$y_3 = 6a_{\mathbf{v}_2} + 2a_{\mathbf{v}_3} + 6a_{\mathbf{v}_6} = 6/160$$
$$y_4 = 8a_{\mathbf{v}_3} + 6a_{\mathbf{v}_5} = 0.$$

The solution is $a_{\mathbf{v}_1} = a_{\mathbf{v}_2} = a_{\mathbf{v}_3} = a_{\mathbf{v}_5} = 0$ and $a_{\mathbf{v}_4} = a_{\mathbf{v}_6} = 1/160$. Thus the D-optimal design consists of the 64 triples with difference vector $\mathbf{v}_4 = (2, 2, 2)$ and the 96 triples with difference vector $\mathbf{v}_6 = (2, 3, 3)$.

However for some values of m and k, solutions to the y_i equations do not exist. For example, when $m = 3$ and $k \equiv 1 \pmod 4$ no solution exists and the following example illustrates this case.

Example 5. If we let $m = 3$ and $k = 5$, then the D-optimal design given by Theorem 2 is

$$y_3 = \frac{3 \times 2}{2^5}\binom{5}{3}^{-1}, \quad y_1 = y_2 = y_4 = y_5 = 0.$$

This means that triples with difference vector $(3, 3, 3)$ are required. However, the only possible difference vectors are $\mathbf{v}_1 = (1, 1, 2)$, $\mathbf{v}_2 = (1, 2, 3)$, $\mathbf{v}_3 = (1, 3, 4)$, $\mathbf{v}_4 = (1, 4, 5)$, $\mathbf{v}_5 = (2, 2, 2)$, $\mathbf{v}_6 = (2, 2, 4)$, $\mathbf{v}_7 = (2, 3, 3)$, $\mathbf{v}_8 = (2, 3, 5)$, $\mathbf{v}_9 = (2, 4, 4)$, and $\mathbf{v}_{10} = (3, 3, 4)$, so no triple with the difference vector $(3, 3, 3)$ exists. For $m = 3$ and $k = 5$ using the techniques in Street et al. (2001), it can easily be shown that the optimal design consists of the 960 triples with difference vector $\mathbf{v}_7 = (2, 3, 3)$ and the 480 triples with difference vector $\mathbf{v}_{10} = (3, 3, 4)$, where $y_2 = 1/1440$, $y_3 = 9/1440$, $y_4 = 3/1440$ and $y_1 = y_5 = 0$. Thus $a_{\mathbf{v}_7} = a_{\mathbf{v}_{10}} = 1/1440$, and $a_{\mathbf{v}_j} = 0$ for $j = 1, 2, 3, 4, 5, 6, 8, 9$.

As Example 5 shows, the optimal designs derived in this and the preceding sections can become very large as the number of attributes increases. The question of how large the number of choice sets can be has been considered by various authors (see Louviere et al. (2002) for a summary). Choice experiments with up to 128 choice experiments have been shown to be equally effective in parameter estimation. In the next section we investigate the D-efficiency of small designs obtained from a generalization of the constant difference construction in Street and Burgess (to appear). For designs that estimate main effects only, the D-efficiency of the proposed design is defined to be the kth root of the ratio of the determinant of the information matrix of the proposed design to that for the optimal design. For designs that estimate main effects and two-factor interactions, the *D*-efficiency is the $(k + k(k-1)/2)$th root of this ratio. Hence the designs derived in Secs. 5 and 6 are 100% efficient.

7. SMALL OPTIMAL AND NEAR-OPTIMAL DESIGNS

In this section we give constructions for small optimal and near-optimal designs for choice sets of size m. The results are an extension of results in Street and Burgess (to appear) who give constructions for optimal and near-optimal designs for testing main effects only, and for main effects and two-factor interactions. The constructions they give start with a fraction of resolution 3 (for testing main effects only) or resolution 5 (for testing main effects and two-factor interactions). Pairs are formed by adding one, or more generators, to the treatments in the fraction, where the addition is performed component-wise modulo 2. Each generator gives rise to a set of pairs.

When testing for main effects only, for each attribute there must be at least one generator with a 1 in the corresponding position. Using one generator consisting of all 1's, which is equivalent to using the resolution 3 fraction and its foldover, results in a *D*-optimal design in a minimum number of pairs, in which all main effects can be estimated.

To estimate main effects and two-factor interactions, Street and Burgess (to appear) give a construction that requires a set of generators, which satisfy two conditions:

1. For each attribute there must be at least one generator with a 1 in the corresponding position (to estimate main effects).
2. For any two attributes there must be at least one generator in which the corresponding positions have a 0 and a 1 (to estimate the two-factor interactions).

The next result gives generators for small, optimal designs for estimating main effects.

Theorem 3. *Let F be a fractional factorial design of resolution at least* 3. *Let* $G = (\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_m)$ *where* $\mathbf{g}_1 = \mathbf{0}$, *be binary k-tuples, which we will call generators. Let* $\mathbf{v} = (d_1, d_2, \ldots, d_{m(m-1)/2})$ *be the difference vector consisting of all the pairwise differences between the generators in G, where*

$$\sum_{i=1}^{m(m-1)/2} d_i = \begin{cases} (m^2-1)k/4, & m \text{ odd}, \\ m^2k/4, & m \text{ even}. \end{cases}$$

Then the choice sets given by $(F, F+\mathbf{g}_2, \ldots, F+\mathbf{g}_m)$, *where the addition is done component-wise modulo* 2, *are optimal for estimating main effects only*.

Proof. Let $|F| = 2^\ell$ and let B_F be the unnormalized contrast matrix for F. Thus $B_F B_F' = 2^\ell I_k$. Let $D_{\mathbf{g}_i}$ be the diagonal matrix with a 1 in position j if the jth entry of $\mathbf{g}_i$ is 0 and a -1 in position j if the jth entry of $\mathbf{g}_i$ is 1. Then $D_{\mathbf{g}_i} B_F$ is the unnormalized contrast matrix for $F+\mathbf{g}_i$.

We construct the choice sets from F by adding the generators to F to get $(F, F+\mathbf{g}_2, \ldots, F+\mathbf{g}_m)$, where each row represents a choice set. Let p be the number of choice sets of size m and let n_{uv} be the number of choice sets that contain the pair of treatments u, v. Let $n_{uv,cd}$ be the number of choice sets which contain u and v in columns c and d (unordered). Then $n_{uv} = \sum_{c,d} n_{uv,cd}$ and we define the diagonal entries to be $n_{uu} = -\sum_{v \neq u} n_{uv,cd}$. Let $H = (-n_{uv})$ and $H_{cd} = (-n_{uv,cd})$, where the rows and columns of H_{cd} are labelled by all the treatments that appear in the choice sets. Thus some of the rows and columns in H_{cd} will be 0 if they correspond to a treatment that does not appear in either column c or d. Then $H = \sum_{c,d} H_{cd}$. Then $\Lambda = H/m^2 p$.

Consider two columns, c and d, say. Then we know that column c contains $F+\mathbf{g}_c$ and d contains $F+\mathbf{g}_d$. If $F+\mathbf{g}_c = F+\mathbf{g}_d$ then $\mathbf{g}_c + \mathbf{g}_d \in F$ and we can write F as $F_1 \cup (F_1 + \mathbf{g}_c + \mathbf{g}_d)$. Then using this order for the elements of F we have

$$H_{cd} = \begin{bmatrix} I_{2^{\ell-1}} & -I_{2^{\ell-1}} \\ -I_{2^{\ell-1}} & I_{2^{\ell-1}} \end{bmatrix}$$

and we can write B_F as $B_F = \begin{bmatrix} B_{F_1} & D_{\mathbf{g}_c+\mathbf{g}_d} B_{F_1} \end{bmatrix}$. Now $B_{F_1} B_{F_1}' = 2^{\ell-1} I_k$, so

$$\begin{aligned} B_F H_{cd} B_F' &= B_{F_1} B_{F_1}' - D_{\mathbf{g}_c+\mathbf{g}_d} B_{F_1} B_{F_1}' \\ &\quad + D_{\mathbf{g}_c+\mathbf{g}_d} B_{F_1} B_{F_1}' D_{\mathbf{g}_c+\mathbf{g}_d} - B_{F_1} B_{F_1}' D_{\mathbf{g}_c+\mathbf{g}_d} \\ &= 2^\ell (I_k - D_{\mathbf{g}_c+\mathbf{g}_d}). \end{aligned}$$

If $\mathbf{g}_c + \mathbf{g}_d \notin F$ then the same argument establishes that

$$B_F H_{cd} B_F' = 2^{\ell+1} (I_k - D_{\mathbf{g}_c+\mathbf{g}_d}).$$

Now we know that

$$\sum_{i=1}^{m(m-1)/2} d_i = \begin{cases} (m^2-1)k/4, & m \text{ odd}, \\ m^2 k/4, & m \text{ even}. \end{cases}$$

Consider the contribution to the $\sum_i d_i$ in a choice set from just one attribute. The attribute will have x 0's and $(m-x)$ 1's to give a total of $x(m-x)$. This contribution is maximized by having $m/2$ 1's and $m/2$ 0's for m even, and $(m-1)/2$ 0's and $(m+1)/2$ 1's (or vice versa) for m odd. Thus the maximum contribution to $\sum_i d_i$ from any one attribute is $m^2/4$ for m even, and $(m^2-1)/4$ for m odd. Thus we see that each attribute must contribute exactly this amount to $\sum_i d_i$ for the optimal designs.

Suppose $m = 2s+1$. Now $\sum_{c,d} D_{\mathbf{g}_c+\mathbf{g}_d}$ is summing $\binom{s}{2} + \binom{s+1}{2}$ entries of 1 in each diagonal position and $s(s+1)$ entries of -1 in each diagonal position. Thus

$$\sum_{c,d} D_{\mathbf{g}_c+\mathbf{g}_d} = \left(\binom{s}{2} + \binom{s+1}{2} - s(s+1)\right) I_k = -\frac{m-1}{2}\ I_k$$

if m is odd. Similarly, if m is even

$$\sum_{c,d} D_{\mathbf{g}_c+\mathbf{g}_d} = -\frac{m}{2}\ I_k.$$

Suppose that none of the $\mathbf{g}_i$ are in F and that $\mathbf{g}_c + \mathbf{g}_d \notin F$ for any pair c and d. Then

$$\begin{aligned}
&\left[\, B_F\ D_{\mathbf{g}_2} B_F \ldots D_{\mathbf{g}_c+\mathbf{g}_d} B_F \,\right] \sum_{c,d} H_{cd} \begin{bmatrix} B'_F \\ B'_F D_{\mathbf{g}_2} \\ \vdots \\ B'_F D_{\mathbf{g}_c+\mathbf{g}_d} \end{bmatrix} \\
&\quad = \sum_{c,d} (2^{\ell+1}(I_k - D_{\mathbf{g}_c+\mathbf{g}_d})) \\
&\quad = \begin{cases} 2^{\ell+1}\left(\binom{m}{2} + \dfrac{m-1}{2}\right) I_k & m \text{ odd,} \\ 2^{\ell+1}\left(\binom{m}{2} + \dfrac{m}{2}\right) I_k & m \text{ even} \end{cases} \\
&\quad = \begin{cases} 2^{\ell+1}\,\frac{m^2-1}{2} I_k & m \text{ odd,} \\ 2^{\ell+1}\,\frac{m^2}{2} I_k & m \text{ even.} \end{cases}
\end{aligned}$$

To get the C matrix we must normalize B_F and divide the entries in H by $m^2 p$. Recalling that $p = 2^\ell$ in this case (since each choice set is unique), we have

$$C = \begin{cases} \dfrac{m^2-1}{m^2 2^k} I_k & m \text{ odd,} \\ \dfrac{1}{2^k} I_k & m \text{ even} \end{cases}$$

as required.

Now suppose that for at least one pair of columns the number of pairs is $2^{\ell-1}$ (that is, $F + \mathbf{g}_c = F + \mathbf{g}_d$ for some c, d). Then we saw that

$$B_F H_{cd} B_F' = 2^{\ell}(I_k - D_{\mathbf{g}_c+\mathbf{g}_d}).$$

But although this works out for the pairs, once the pairs are considered as part of the larger choice sets then the pairs will in fact appear twice and so we need to use

$$2B_F H_{cd} B_F' = 2^{\ell+1}(I_k - D_{\mathbf{g}_c+\mathbf{g}_d})$$

when evaluating C. Thus the proof from before can be used.

The only situation that is not covered by the above proof is when m is a power of 2. In that case the optimal set of generators must form a subgroup and the set of choice sets are this subgroup and its distinct cosets formed by adding elements of F. Making this observation, a straight-forward modification of the proof above establishes the result.

Example 6. Let $m = 5$ and $k = 9$. To obtain an optimal design for estimating main effects, we require a fraction F of the 2^9 factorial which has resolution at least 3. The 16 treatments given in the first column of Table 1 are a 1/32 fraction of resolution 3 with defining contrast $I \equiv BCE \equiv CDF \equiv ACG \equiv ABH \equiv ADJ$. To obtain the choice sets we need $m = 5$ generators $G = (\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3, \mathbf{g}_4, \mathbf{g}_5)$ where $\mathbf{g}_1 = \mathbf{0}$, so that the

Table 1. Optimal choice sets for estimating main effects only for $m = 5$ and $k = 9$.

(000000000, 000000111, 111111000, 000111111, 111111111)
(000101001, 000101110, 111010001, 000010110, 111010110)
(001011100, 001011011, 110100100, 001100011, 110100011)
(001110101, 001110010, 110001101, 001001010, 110001010)
(010010010, 010010101, 101101010, 010101101, 101101101)
(010111011, 010111100, 101000011, 010000100, 101000100)
(011001110, 011001001, 100110110, 011110001, 100110001)
(011100111, 011100000, 100011111, 011011000, 100011000)
(100000111, 100000000, 011111111, 100111000, 011111000)
(100101110, 100101001, 011010110, 100010001, 011010001)
(101011011, 101011100, 010100011, 101100100, 010100100)
(101110010, 101110101, 010001010, 101001101, 010001101)
(110010101, 110010010, 001101101, 110101010, 001101010)
(110111100, 110111011, 001000100, 110000011, 001000011)
(111001001, 111001110, 000110001, 111110110, 000110110)
(111100000, 111100111, 000011000, 111011111, 000011111)

differences in the difference vector sum to $(m^2-1)k/4 = 6k = 54$. One set of generators that satisfies this condition is

$$G = (000000000,\ 000000111,\ 111111000,\ 000111111,\ 111111111)$$

which has the difference vector $(3,3,3,3,6,6,6,6,9,9)$. The 16 choice sets are $(F,\ F+\mathbf{g}_2,\ F+\mathbf{g}_3,\ F+\mathbf{g}_4,\ F+\mathbf{g}_5)$ where the addition is done component-wise modulo 2. The choice sets are given in Table 1, where the choice sets are represented by the rows. This design is 100% efficient.

To estimate main effects and two-factor interactions in paired comparisons, a construction is given in Street and Burgess (to appear) and starts with a resolution 5 (or more) fraction of the the complete 2^k factorial. Sets of generators are added to this fraction to obtain near-optimal pairs. This method can easily be extended to obtain near-optimal choice sets of size m.

Let $G_j = (\mathbf{g}_{1j}, \mathbf{g}_{2j}, \ldots, \mathbf{g}_{mj})$, where $\mathbf{g}_1 = \mathbf{0}$, be binary k-tuples which we will call generators. Let $\mathbf{v}_j = (d_{1j}, d_{2j}, \ldots, d_{m(m-1)/2,j})$ be the difference vector consisting of all the pairwise differences between the generators in G_j. The possible $\mathbf{v}_j$ vectors are those difference vectors determined in Sec. 6 for an optimal design for the particular values of m and k. Thus the G_j are not unique and several different near-optimal designs are possible. The construction of the choice sets is as follows:

1. Start with a resolution 5 (or more) fraction F of the complete 2^k factorial design. Let F have 2^ℓ treatments.
2. Add generator G_1 to F where the addition is done component-wise modulo 2, to form 2^ℓ choice sets of size m.
3. Repeat step 2, to form another 2^ℓ choice sets of size m, until all main effects and two-factor interactions can be estimated. That is, for each attribute there must be at least one generator with a 1 in the corresponding position (to estimate main effects) and for any two attributes, there must be at least one generator in which the corresponding positions have a 0 and a 1 (to estimate the two-factor interactions).

In many cases step 2 is only repeated once, so the number of choice sets usually required for a near-optimal design is $2^{\ell+1}$.

Example 7. Let $m = 3$ and $k = 4$. There are no fractions of the 2^4 factorial which are resolution 5, so we must use the 16 treatments from the complete factorial. These treatments are given in the first column of Table 2. In Example 4 the difference vectors for the

Table 2. Near-optimal choice sets for estimating main effects and two-factor interactions for $m = 3$ and $k = 4$.

(0 0 0 0, 1 1 0 0, 0 1 1 0)	(0 0 0 0, 1 1 0 0, 0 1 1 1)
(0 0 0 1, 1 1 0 1, 0 1 1 1)	(0 0 0 1, 1 1 0 1, 0 1 1 0)
(0 0 1 0, 1 1 1 0, 0 1 0 0)	(0 0 1 0, 1 1 1 0, 0 1 0 1)
(0 0 1 1, 1 1 1 1, 0 1 0 1)	(0 0 1 1, 1 1 1 1, 0 1 0 0)
(0 1 0 0, 1 0 0 0, 0 0 1 0)	(0 1 0 0, 1 0 0 0, 0 0 1 1)
(0 1 0 1, 1 0 0 1, 0 0 1 1)	(0 1 0 1, 1 0 0 1, 0 0 1 0)
(0 1 1 0, 1 0 1 0, 0 0 0 0)	(0 1 1 0, 1 0 1 0, 0 0 0 1)
(0 1 1 1, 1 0 1 1, 0 0 0 1)	(0 1 1 1, 1 0 1 1, 0 0 0 0)
(1 0 0 0, 0 1 0 0, 1 1 1 0)	(1 0 0 0, 0 1 0 0, 1 1 1 1)
(1 0 0 1, 0 1 0 1, 1 1 1 1)	(1 0 0 1, 0 1 0 1, 1 1 1 0)
(1 0 1 0, 0 1 1 0, 1 1 0 0)	(1 0 1 0, 0 1 1 0, 1 1 0 1)
(1 0 1 1, 0 1 1 1, 1 1 0 1)	(1 0 1 1, 0 1 1 1, 1 1 0 0)
(1 1 0 0, 0 0 0 0, 1 0 1 0)	(1 1 0 0, 0 0 0 0, 1 0 1 1)
(1 1 0 1, 0 0 0 1, 1 0 1 1)	(1 1 0 1, 0 0 0 1, 1 0 1 0)
(1 1 1 0, 0 0 1 0, 1 0 0 0)	(1 1 1 0, 0 0 1 0, 1 0 0 1)
(1 1 1 1, 0 0 1 1, 1 0 0 1)	(1 1 1 1, 0 0 1 1, 1 0 0 0)

optimal designs are $(2, 2, 2)$ and $(2, 3, 3)$, so $\mathbf{g}_{2j}$ and $\mathbf{g}_{3j}$ must be chosen so that the difference vector for G_j is either $(2, 2, 2)$ or $(2, 3, 3)$. We choose $G_1 = (0000, 1100, 0110)$ with $\mathbf{v}_1 = (2, 2, 2)$ and form $2^\ell = 16$ triples by adding G_1 to F. From these we can estimate all the main effects and two-factor interactions except the main effect of the fourth attribute. So we repeat step 2 adding $G_2 = (0000, 1100, 0111)$, where $\mathbf{v}_2 = (2, 3, 3)$, to F to get an additional 16 triples. All main effects and two-factor interactions can now be estimated so we have no need to generate any more triples. The 32 triples shown in Table 2 form a design that is 96.73% efficient and is thus near-optimal. This design is much smaller than the optimal design in Example 4 which consists of 160 triples.

8. DISCUSSION

There are two standard ways to design choice sets of size m with each choice having k attributes.

The first of these is to construct a resolution 3 fractional factorial design with km factors; see, for example, Option 4 in Appendix A5 of Louviere et al. (2000). In Kuhfeld (2000) a thorough discussion of SAS macros which perform this construction is given.

Unfortunately these designs are rarely able to estimate main effects orthogonally in a forced choice setting since each choice comes from only k of the attributes of the original resolution 3 design and no requirements are placed on subsets of the attributes. The designs are also larger than the ones we have given here. For instance, the design of Example 6 would require at least 48 choice sets compared with the 16 we have given in Table 2. The Kuhfeld macros quote the OLS optimality values for the designs that they generate. Our methods starts with a design that is D-optimal in the OLS setting. In comparing designs in this article we use the optimal values assuming the treatment combinations used over all the choice sets allow for an orthogonal matrix of main effects contrasts to be calculated, that an MNL model is used and that the null hypothesis is true. As the "2^{km} approach" does not give an orthogonal matrix of main effects in any of the instances that we have tried, we have been unable to calculate a D-optimal value for the setting of Example 6.

The other technique that is used to generate designs for this setting depends on the MNL being the correct representation of the choice process. In that case take m (different) resolution 3 designs on k attributes and randomly choose, without replacement, one profile from each resolution 3 design to form a choice set. This process is described in Option 1 of Appendix A5 of Louviere et al. (2000). Often this process results in designs with the same problems described in the previous paragraph. For the setting of Example 6 we did get one design with an orthogonal matrix for main effects. The design had 16 choice sets and had an information matrix in which only five pairs of main effects were independently estimated. The design was 80.42% efficient.

Thus we believe that the methods described in this article give small, optimal designs that are at least as easy to construct as currently used designs.

ACKNOWLEDGMENT

This research was supported by the Australian Research Council by grant number A79906045.

REFERENCES

Bradley, R. A. (1955). Rank analysis of incomplete block designs. III. Some large-sample results on estimation and power for a method of paired comparisons. *Biometrika* 42:450–470.

Bradley, R. A., Terry, M. E. (1952). Rank analysis of incomplete block designs. I. The method of paired comparisons. *Biometrika* 39:324–345.

Bunch, D. S., Louviere, J. J., Anderson, D. A. (1996). A comparison of experimental design strategies for choice-based conjoint analysis with generic-attribute multinomial logit models. Working Paper, Graduate School of Management, UC Davis.

El-Helbawy, A. T., Ahmed, E. A. (1984). Optimal design results for 2^n factorial paired comparison experiments. *Communications in Statistics—Theory and Methods* 13:2827–2845.

El-Helbawy, A. T., Ahmed, E. A., Alharbey, A. H. (1994). Optimal designs for asymmetrical factorial paired comparison experiments. *Communications in Statistics—Simulation* 23:663–681.

Kuhfeld, W. F. (2000). Multinomial Logit, Discrete Choice Modelling, SAS Institute. Available at http://ftp.sas.com/techsup/download/technote/ts621.pdf.

Louviere, J. J., Hensher, D. A., Swait, J. D. (2000). *Stated Choice Models: Analysis and Application*. Cambridge: Cambridge University Press.

Louviere, J. J., Street, D. J., Carson, R., Ainslie, A., Deshazo, J. R., Cameron, T., Hensher, D., Kohn, R., Marley, T. (2002). Dissecting the random component of utility. *Marketing Letters* 13:177–193.

Moore, B. J. (2001). Personal communication.

Pendergrass, R. N., Bradley, R. A. (1960). Ranking in triple comparisons. In: Olkin, I., Ghurye, S. G., Hoeffding, W., Madow, W. G., Mann, H. B., eds. *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. Stanford: Stanford University Press, 331–351.

Street, D. J., Bunch, D. S., Moore, B. (2001). Optimal designs for 2^k paired comparison experiments. *Communications in Statistics—Theory and Methods* 30:2149–2171.

Street, D. J., Burgess, L. (to appear). Optimal and near-optimal pairs for the estimation of effects in 2-level choice experiments. *Journal of Statistical Planning and Inference*.

van Berkum, E. E. M. (1987). Optimal paired comparison designs for factorial and quadratic models. *Journal of Statistical Planning and Inference* 15:265–278.