Optimal Stated Preference Choice Experiments When All Choice Sets Contain A Specific Option

Deborah J. Street [a,*,1], Leonie Burgess [a]

[a]*Department of Mathematical Sciences, University of Technology Sydney, PO Box 123, Broadway, NSW 2007, Australia*

Abstract

Stated preference choice experiments are routinely used in many areas from marketing to medicine. While results on the optimal choice sets to present for the forced choice setting have been determined in a variety of situations, no results have appeared to date on the optimal choice sets to use when either all choice sets are to contain a common base alternative or when all choice sets contain a 'none of these' option. These problems are considered in this paper.

Key words: Paired comparisons, Multiple comparisons, Bradley-Terry model, Multinomial logit model, Fractional factorial designs, Orthogonal main effect plans
1991 MSC: primary 62J15, secondary 62K05

* Corresponding author.
  Email address: Deborah.Street@uts.edu.au (Deborah J. Street).
[1] Phone: 61 2 9514 2251; Fax: 61 2 9514 2260

1 Introduction

Stated choice experiments are widely used in various areas including health research, marketing, transport and public welfare analysis; see Louviere, Hensher and Swait (1) for an introduction to stated choice methods. A *forced choice stated preference* experiment is an experiment in which several choice sets are presented, in turn, to respondents and one of the options must be chosen from each choice set presented.

Suppose that each option in a choice set is described by k attributes and each choice set contains m options. We will assume that the qth attribute has ℓ_q levels, represented by $0, 1, \ldots, \ell_q - 1$ and attributes may have different numbers of levels. Under these circumstances results about the optimal forced choice stated preference experiments for the estimation of main effects for any number of attributes with any number of levels may be found in Burgess and Street (2) and references cited therein. They also make some comments about finding good designs for the estimation of main effects plus two-factor interactions and some near-optimal designs may be found in Burgess and Street (3) and Street and Burgess (4).

Scott (5) describes a discrete choice experiment to do investigate attributes that have an effect on parent satisfaction with after hours medical care provided to children under 13. He used four attributes to describe each option. The first attribute was *where your child is seen* with levels *emergency centre run by GPs*, *your home* and *hospital accident and emergency department*. The second attribute was *Who sees your child* with levels *a GP from your practice/health centre* and *a GP who does not work at your practice/health centre*. The third attribute was *Time taken between the telephone call and treatment being received* with levels *20 minutes*, *40 minutes*, *60 minutes* and *80 minutes*. The final attribute is the *attentiveness of the doctor* with levels *the doctor seems to listen* and *the doctor does not seem to listen*. These attributes are then combined to give descriptions of possible consultations and respondents are asked which of two consultations they would prefer. A typical choice set appears in Table 1.

It seems that no results have been published to date on the construction of optimal stated preference choice experiments when a common base is to appear in each choice set or when a "none of these" option is to be available in each choice set. Statistically optimal designs for these situations are presented in this paper. (For a discussion of how to design experiments to avoid consumers choosing to defer choice, the "none of these option", see Haaijer, Kamakura and Wedel (6) and references cited therein. For a discussion of respondent efficiency and the role of cognitive complexity in the design of choice experiments readers are referred to De Shazo and Fermo (7) and Arentze, Borgers,

Table 1
One choice set in an experiment

Suppose that you have to consult a GP during the night because your child is sick.

Which of these two consultation would you prefer?

	Consultation A	Consultation B
Where your child is seen	Your home	Emergency centre run by GPs
Who your child sees	GP from your practice	GP from your practice
Time taken to receive treatment	80 minutes	20 minutes
Attentiveness	Seems to listen	Does not seem to listen

Timmeramns and Del Mistro (8) and references cited therein.)

2   Common Base Option

If a particular combination of attribute levels appears in all choice sets then this combination is called the *common base option*. It may represent the current situation or the current treatment for a particular health condition (examples in Ryan and Hughes (9), Ryan and Farrar (10), Longworth, Ratcliffe and Boulton (11)) or the common base may be randomly chosen from the main effects plan and have all the other scenarios from the plan compared to it pairwise (examples in Ryan (12), Ryan, McIntosh, Dean and Old (13), Scott (5)).

To ensure that the matrix of contrasts for main effects is unambiguously defined we will insist that the set of all treatment combinations that appear in the choice experiment forms a fractional factorial design of resolution 3. Lewis and John (14) discuss the appropriate structure of contrasts when there is not equal replication of levels.

Our first optimality result gives the determinant of the information matrix for any resolution 3 fractional factorial design in which any one of the treatment combinations may be used as the common base. First we need the following result about the main effects contrast matrix B.

Theorem 1 *Let B be a normalised main effects contrast matrix corresponding to all the attributes in a $\ell_1 \times \ell_2 \times \ldots \ell_k // t$ resolution 3 fractional factorial design. Then the diagonal entries of $B'B$ are all equal to $\sum_{q=1}^{k}(\ell_q - 1)/t$.*

Proof. Suppose that Q_ℓ is an orthonormal matrix of order ℓ with all the entries of the first row equal. Then $Q_\ell Q'_\ell = Q'_\ell Q_\ell = I_\ell$. Let B_ℓ be the $(\ell-1) \times \ell$ matrix obtained from Q_ℓ by removing the first row. Then

$$Q'_\ell Q_\ell = \tfrac{1}{\ell} \mathbf{j}'_\ell \mathbf{j}_\ell + B'_\ell B_\ell = \tfrac{1}{\ell} J_\ell + B'_\ell B_\ell = I_\ell$$

where $\mathbf{j}_\ell$ is a $1 \times \ell$ vector of ones and J_ℓ is an $\ell \times \ell$ matrix of ones. Thus $B'_\ell B_\ell = I_\ell - 1/\ell J_\ell$. Note that B_ℓ is a main effects matrix for an attribute with ℓ levels.

Now consider a main effects matrix, B say, corresponding to all the attributes in a $\ell_1 \times \ell_2 \times \cdots \times \ell_k // t$ resolution 3 fractional factorial design. For the qth attribute the matrix B_{ℓ_q}, possibly with columns permuted, appears t/ℓ_q times in the main effects matrix for the fractional factorial design. The entries in each such B_{ℓ_q} will have been normalised by division by $\sqrt{t/\ell_q}$.

The diagonal entries of $B'B$ equal the sum of the diagonal entries in the $B'_{\ell_q} B_{\ell_q}$ which appears in B, suitably normalised. Thus all the diagonal entries are equal to

$$\sum_{q=1}^{k} \left((1 - \tfrac{1}{\ell_q}) \tfrac{1}{t/\ell_q} \right) = \sum_{q=1}^{k} (\ell_q - 1)/t,$$

as required.

Now we can explicitly evaluate the determinant of the information matrix, which is given by $C = B \Lambda B'$, where Λ is the matrix of second derivatives of the likelihood function for the multinomial logit (MNL) model (see El-Helbawy and Bradley (15) or Burgess and Street (2)). We will let $w = \sum_{q=1}^{k} (\ell_q - 1)$, the total degrees of freedom for main effects, and $L = \prod_{q=1}^{k} \ell_q$.

From Theorem 2 in Burgess and Street (2), we know that the maximum possible determinant for the information matrix, when choice sets are of size 2, is given by

$$\prod_{q=1}^{k} \left[\frac{1}{2(\ell_q - 1)L/\ell_q} \right]^{\ell_q - 1} = optdet, \quad \text{say.}$$

This is for a forced choice stated preference experiment with no restrictions on the options that can appear in each choice set.

Theorem 2 *Let F be a resolution 3 fractional factorial design with equal replication of levels and t treatment combinations. Choose any treatment combination in F to be the common base. Then*

$$\det(C) = \frac{1+w}{(4(t-1))^w} \left(\frac{t}{L} \right)^w.$$

The efficiency of the design, relative to the optimal forced choice stated pref-

4

erence experiment with choice sets of size 2, is given by

$$\left(\frac{\det(C)}{optdet} \right)^{\frac{1}{w}} \quad .$$

Proof. Without loss of generality, order the treatment combinations so that the first treatment combination is the common base and let B be the normalised main effects contrast matrix. Then $BB' = I_w$ and

$$4(t-1)\Lambda = \begin{bmatrix} (t-1) & -1 & -1 & \dots & -1 \\ -1 & 1 & 0 & \dots & 0 \\ -1 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ -1 & 0 & 0 & \dots & 1 \end{bmatrix}.$$

Then $C = tB\Lambda B'/L$, since we have normalised B without allowing for the fact that it is derived only from the treatment combinations in a fraction, so

$$C = \tfrac{t}{4(t-1)L}(BB' + t\mathbf{b}_1\mathbf{b}_1')$$

where $\mathbf{b}_1$ is the first column of B. Since $\mathbf{b}_1\mathbf{b}_1'$ is the first diagonal entry of $B'B$, we see that

$$\det(C) = \left(\tfrac{t}{4(t-1)L}\right)^w \det(I_w + t\mathbf{b}_1\mathbf{b}_1') = (1+w)\left(\tfrac{t}{4(t-1)L}\right)^w$$

by Theorem 1. (We need the result that for any vector $\mathbf{a}$ it is true that $\det(I_n + \mathbf{a}\mathbf{a}') = 1 + \sum_i a_i^2$.)

By definition, the efficiency relative to the unrestricted forced choice experiment is $(\det(C)/optdet)^{1/w}$, as required.

Using Theorem 2, we can see that from the point of view of statistical efficiency all resolution 3 designs with the same number of treatment combinations are equally good for a particular number of attributes with given number of levels. It is also immaterial which of the treatment combinations is used as the common base.

Example 1 *Consider an experiment where there are four attributes describing each option. Three of the attributes have 2 levels and one has 4 levels. Then using any $2 \times 2 \times 2 \times 4//16$ and using any treatment combination as a common base gives a design which is 45.17% efficient relative to a forced*

choice stated preference design. Using a $2 \times 2 \times 2 \times 4//8$ gives an efficiency of 48.40%, while using the complete factorial gives an efficiency of 43.71%.

In Scott(5) a similar experiment was carried out with four attributes, two with 2 levels, one with 3 levels and one with 4 levels, and using 16 choice sets. In this case the efficiency depends on the treatment combination chosen to be the common base. If the common base has the level of the 3 level attribute which appears 8 times then the efficiency is 44.49% and if the common base has either of the other levels of the 3 level attribute then the efficiency is 46.12%.

If we want to use this approach to estimate main effects plus two-factor interactions then we need to start with a resolution 5 fractional factorial design. Unfortunately in this setting we can only know that all effects are estimable and compare particular designs. We cannot calculate the efficiency relative to the optimal design since only a general expression for $\det(C)$ is available (see Burgess and Street (2)).

3 The 'None of These' Option

In this section we consider what happens when we adjoin a "none of these" option to each choice set in a stated preference choice experiment. It turns out that there is a simple relationship between the matrices for a forced choice stated preference experiment and those from the the same choice sets with a "none of these" option adjoined to each choice set.

We let B_f be the contrast matrix for the forced choice experiment, Λ_f be the matrix of second derivatives of the likelihood function and let $C_f = B_f \Lambda_f B'_f$ be the information matrix for the forced choice experiment. We assume that B_f is $c \times L$ so there are c contrasts of interest and note that Λ_f will contain rows and columns of 0s if not all treatment combinations appear in the choice experiment. We will use B_n, Λ_n and $C_n = B_n \Lambda_n B'_n$ for the corresponding matrices when a "none of these" option has been adjoined. We assume that each choice set in the forced choice experiment has m options in it and that there are p such choice sets.

The next result establishes the relationship between B_f and B_n and between Λ_f and Λ_n.

Theorem 3 *(1) Let $d^2 = L(L+1)$. Then*

$$B_n = \begin{bmatrix} B_f & \mathbf{0}'_c \\ \frac{1}{d}\mathbf{j}_L & \frac{-L}{d} \end{bmatrix}$$

where $\mathbf{0}_c$ is a $1 \times c$ vector of zeroes and "none" is the final treatment. Thus $B_n B'_n = I_{c+1}$.

(2) Let r_i be the number of times the ith treatment combination appears in the stated preference experiment and $\mathbf{r} = (r_1, \ldots, r_L)$. Let D be a diagonal matrix with these replication numbers on the diagonal. Then

$$\Lambda_n = \frac{1}{(m+1)^2 p} \begin{bmatrix} m^2 p \Lambda_f + D & -\mathbf{r}' \\ -\mathbf{r} & mp \end{bmatrix}$$

Proof.

(1) The only additional contrast is between "none" and some treatment combination. The divisor d ensures that the contrast is of unit length.

(2) The "none" option will appear with each treatment combination as many times as that treatment combination appears in the design so $r_{none} = mp = \sum_i r_i$. Again only one additional row and column need to be adjoined to Λ.

Now we are in a position to evaluate C_n.

Theorem 4 *Consider a forced choice stated preference experiment in which a "none" option has been adjoined to each choice set. Then*

$$C_n = \frac{1}{(m+1)^2 p} \begin{bmatrix} m^2 p C_f + B_f D B'_f & \frac{L+1}{d} B_f \mathbf{r}' \\ \frac{L+1}{d} \mathbf{r} B'_f & \frac{mp(L+1)}{L} \end{bmatrix}$$

Proof. This result follows directly from the definition, noting that $\mathbf{j}_L \mathbf{r}' = mp$, $\Lambda_f \mathbf{j}'_L = 0$ and $\mathbf{j}_L D = \mathbf{r}$.

Recall that the treatment combinations in an $\ell_1 \times \ell_2 \times \ldots \times \ell_k$ complete factorial design can be thought of the elements in an Abelian group of the same order. Thus we can represent the treatment combinations as the elements of a subgroup (of the Abelian group) and the elements of the cosets of that subgroup. Within each coset each level of each attribute appears equally often. (It may be more familiar to think of the subgroup as the principal block and the cosets as the other blocks in a blocked factorial design.) Using this observation we can establish the following result.

Theorem 5 *Suppose that all the treatment combinations in each coset (of any fixed but arbitrary subgroup) appear the same number of times, perhaps 0*

times, in the choice experiment. Then

$$C_n = \frac{m}{(m+1)^2} \begin{bmatrix} mC_f + \frac{1}{L}I_c & \mathbf{0}'_c \\ \mathbf{0}_c & \frac{L+1}{L} \end{bmatrix}$$

Proof. Without loss of generality we can order the treatment combinations so that all treatment combinations from the subgroup appear first, then all of those from some coset and so on. We assume that there are s distinct cosets (including the subgroup) and that t_i is the number of times each treatment in the ith coset is replicated. Now $\sum_i r_i = mp$ and each coset has L/s treatment combinations in it so $\sum_i t_i = (mps/L)$.

We can write

$$B_n = \begin{bmatrix} B_1 \ B_2 \ldots B_s & \mathbf{0}'_w \\ \frac{1}{d}\mathbf{j}_L & \frac{-L}{d} \end{bmatrix}$$

where B_i is the B matrix for the ith coset. Since $B_f B'_f = I_c = \sum_{i=1}^{s} B_i B'_i$ we have that $B_i B'_i = (1/s)I_c$ for $i = 1, 2, \ldots, s$.

Since each treatment combination in each coset appears the same number of times, $B_f \mathbf{r}' = \mathbf{0}_c$. Also $B_f D B'_f = \sum_{i=1}^{s} t_i B_i B'_i = \sum_{i=1}^{s} (t_i/s)I_c$. The result follows.

Although this result may appear to be rather restrictive at first glance, in fact most of the constructions for optimal design presented in the literature give designs which satisfy the conditions of the Theorem 5. For example optimal designs for estimating main effects, given in Street and Burgess (4), have this structure. These designs have

$$C_f = \frac{2}{m^2 L}\text{BlkDiag}\left(\frac{S_1\ell_1}{\ell_1-1}I_{\ell_1-1}, \frac{S_2\ell_2}{\ell_2-1}I_{\ell_2-1}, \ldots, \frac{S_k\ell_k}{\ell_k-1}I_{\ell_k-1}\right)$$

where

$$S_q = \begin{cases} (m^2-1)/4 & \ell_q = 2, \ m \text{ odd}, \\ m^2/4 & \ell_q = 2, \ m \text{ even}, \\ (m^2 - (\ell_q x^2 + 2xy + y))/2 & 2 < \ell_q \le m, \\ m(m-1)/2 & \ell_q \ge m \end{cases}$$

and positive integers x and y satisfy the equation $m = \ell_q x + y$ for $0 \le y < \ell_q$.

So

$$\det(C_f) = \prod_{q=1}^{k} \left[\frac{2\ell_q S_q}{m^2(\ell_q - 1)L} \right]^{\ell_q - 1}.$$

If we adjoin a 'none' option to each choice set then

$$C_n = \tfrac{1}{L(m+1)^2} \text{BlkDiag}\left(\tfrac{2S_1\ell_1 + (\ell_1 - 1)}{\ell_1 - 1} I_{\ell_1 - 1}, \ldots, \tfrac{2S_k\ell_k + (\ell_k - 1)}{\ell_k - 1} I_{\ell_k - 1}, m(L + 1)I_1 \right).$$

Thus

$$\det(C_n) = \prod_{q=1}^{k} \left[\frac{2\ell_q S_q + (\ell_q - 1)}{(m + 1)^2(\ell_q - 1)L} \right]^{\ell_q - 1} \times \frac{m(L + 1)}{L(m + 1)^2}.$$

Hence the same designs that are optimal for the estimation of main effects in the forced choice setting are optimal for the estimation of main effects when a "none of these" option has been adjoined to each choice set, although the efficiency of the design is reduced from 100% to $m^2\{\prod_{q=1}^{k}[\ell_q S_q + (\ell_q - 1)/\ell_q S_q]\}^{1/w}/(m + 1)^2$.

But adjoining a "none" option does change the properties of the design. If there is no "none" option then the optimal design for main effects can not be used to give any information about two-factor interactions. The inclusion of a "none" option makes it possible to estimate two-factor interactions as well, although the efficiency may not be very high.

Example 2 *Suppose that we have $k = 3$ attributes with $\ell_1 = \ell_2 = 2$ and $\ell_3 = 3$. The optimal design using pairs for estimating main effects has the following 12 choice sets: (000, 111), (001, 112), (002, 110), (010, 101), (011, 102), (012, 100), (100, 011), (101, 012), (102, 010), (110, 001), (111, 002), (112, 000).*

The optimal design using pairs for estimating main effects plus two factor interactions has the following 30 choice sets: (000, 011), (001, 012), (002, 010), (010, 001), (011, 002), (012, 000), (100, 111), (101, 112), (102, 110), (110, 101), (111, 102), (112, 100), (000, 101), (001, 102), (002, 100), (010, 111), (011, 112), (012, 110), (100, 001), (101, 002), (102, 000), (110, 011), (111, 012), (112, 010), (000, 110), (001, 111), (002, 112), (010, 100), (011, 101), (012, 102). The determinant of the corresponding information matrix is 2.60417×10^{-12}.

If we adjoin a "none of these" option to the 12 choice sets which are optimal for estimating main effects then the determinant for the part of the information matrix that corresponds to main effects plus two factor interactions is 7.29363×10^{-14} and so the design is 67.2% efficient for estimating these effects. The same set of 12 choice sets are now only 70.3% efficient for estimating main effects. In this example all effects are independently estimated in

all of the designs discussed.

4   Concluding Remarks

If it is important to have a common base in a forced choice, stated preference experiment then the most statistically efficient design for estimating main effects uses the smallest possible resolution 3 design.

Including a "none of these" option in a set of choice sets that are optimal for main effects reduces their efficiency for the estimation of main effects but does mean that the sets can be used to estimate two factor interactions as well. It does not seem to be possible to quantify these changes in efficiency in general.

Finally note that we have assumed that all components of the main effects are of equal importance. If only the linear component was of interest then it is possible that smaller designs could be used.

References

[1]   Louviere, J.J., Hensher, D.A. and Swait, J.D., 2000. Stated choice methods: analysis and application. Cambridge University Press, Cambridge.

[2]   Burgess, L. and Street, D.J., 2003b. Optimal designs for asymmetric choice experiments. Research Report, Department of Mathematical Sciences, University of Technology, Sydney.

[3]   Burgess, L. and Street, D.J., 2003a. Optimal designs for 2^k choice experiments, Communications in Statistics - Theory and Methods 32, 2185-2206.

[4]   Street, D.J. and Burgess, L. 2004. Optimal and near-optimal pairs for the estimation of effects in 2-level choice experiments. Journal of Statistical Planning and Inference 118, 185-199.

[5]   Scott, A., 2002. Identifying and analysing dominant preferences in discrete choice experiments: An application in health care, Journal of Economic Psychology 23, 383-398.

[6]   Haaijer, R., Kamakura, W. and Wedel, M. (2001). The "no-choice" alternative in conjoint choice experiments, Journal of Market Research 43, 93-106.

[7]   De Shazo, J.R. and Fermo, G., 2002. Designing chioce sets for stated pref-

erence methods: The effects of complexity on choice consistency, Journal of Environmental Economics and Management 44, 123-143.

[8] Arentze, T., Borgers, A., Timmermans, H., and Del Mistro, R., (2003). Transport stated choice responses: Effects of task complexity, presentation format and literacy, Transportation Research 39E, 229-244.

[9] Ryan, M. and Hughes, J., 1997. Using conjoint analysis to assess women's preferences for miscarriage management, Health Economics 6, 261-273.

[10] Ryan, M. and Farrar, S. 2000. Using conjoint analysis to elicit preferences for health care, British Mdeical Journal 320, 1530-1533.

[11] Longworth, L., Ratcliffe, J. and Boulton, M., 2001. Investigating women's preferences for intrapartum care: Home versus hospital births, Health and Social Care in the Community 9, 404-413.

[12] Ryan, M., 1999. Using conjoint analysis to take account of patient preferences and go beyond health outcomes: An application to in vitro fertilisation, Social Science and Mdeicine 48, 535-546.

[13] Ryan, M., McIntosh, E., Dean, T. and Old, P., 2000. Trade-offs between location and waiting times in the prevision of health care: The case of elective surgery on the Isle of Wight, Journal of Public Health Medicine 22, 202-210.

[14] Lewis, S.M. and John, J.A. (1976). Testing main effects in fractions of asymmetrical factorial experiments, Biometrika 63, 678-680.

[15] El-Helbawy, A.T. and Bradley,R.A., (1978). Treatment contrasts in paired comparisons: Large-sample results, applcations and some optimal designs, Journal of the American Statistical Association 73, 831-839.