Technical Report 2015

Optimal Paired Choice Block Designs

Rakhi Singh[1], Ashish Das[2] and Feng-Shun Chai[3]

[1]IITB-Monash Research Academy, IIT Bombay, Mumbai, India
[2]Indian Institute of Technology Bombay, Mumbai, India
[3]Academia Sinica, Taipei, Taiwan

15th September 2015

Indian Institute of Technology Bombay
Powai, Mumbai-400 076, India

Optimal Paired Choice Block Designs

RAKHI SINGH
IITB-Monash Research Academy, Mumbai, India

ASHISH DAS
Indian Institute of Technology Bombay, Mumbai, India

FENG-SHUN CHAI
Academia Sinica, Taipei, Taiwan

Abstract

Choice experiments mirror real world situations closely and helps manufacturers, policy-makers and other researchers in taking business decisions on their product characteristics based on its perceived utility. In a paired choice experiment, several pairs of options are shown to respondents. The respondents are asked to give their preference among the two options for each of the choice pairs shown to them. In order to conduct an experiment, a choice design is customarily used to efficiently estimate the parameters of interest which essentially consists of either the main effects only or the main plus two-factor interaction effects of the attributes. Traditionally, every respondent is shown the same collection of choice pairs under an untenable assumption that the respondents are alike in every respect. Also, as the attributes or the number of levels under each attribute increases, the number of choice pairs in an optimal paired choice design increases rapidly. To address these concerns, under the multinomial logit model or the linear paired comparison model, we first incorporate the respondent effects and then present optimal designs for the parameters of interest. We provide optimal paired choice designs for estimating the main effects for symmetric and asymmetric multi-level attributes with smaller number of choice pairs shown to each respondent. We also provide optimal paired choice designs for estimating the main effects only and the main plus two-factor interaction effects under the main plus two-factor interaction effects model.

1 Introduction

Choice experiments play an instrumental role in marketing, transport, environmental resource economics and public welfare analysis. In a choice experiment, respondents are shown multiple choice sets of options and from each set they choose the preferred option. Considering choice sets of size two and b given respondents, a paired choice experiment consists of showing the same set of N choice pairs to each of the b respondents. The respondents are asked to give their preference among the two options for each of the

N choice pairs shown to them. Thus, the choice experiments comprises of a total of $n = bN$ choice pairs. Though choice designs may contain repeated choice sets, one may prefer that no two choice sets are repeated. Each option in a choice pair is described by a set of k attributes, where for $i = 1, \ldots, k$, the ith attribute has v_i levels, $v_i \geq 2$. We represent the v_i levels by $0, \ldots, v_i - 1$. In a choice experiment, a paired choice design d is an allocation of $n = bN$ choice pairs among b respondents such that each respondent observes N choice pairs. Such paired choice designs are analysed under the multinomial logit model and the linear paired comparison model.

The objective of a choice experiment is to efficiently estimate the parameters of interest which essentially consists of either only the main effects or the main plus two-factor interaction effects of the k attributes. Traditionally, in a choice experiment, each of the b respondents are shown the same collection of N choice pairs under the assumption that the respondents are alike or that showing the same collection of choice pairs to the b respondents would eliminate the respondent variations with respect to extraneous effects. Therefore, while comparing choice designs, it has been sufficient to compare designs with respect to only N choice pairs.

The choice experiments with the inherent premise that the respondents are alike or that the same set of N choice pairs needs to be shown to the b respondents is not quite practical since respondents being a random sample from a population is more likely to be heterogeneous. Also as noted in [5], heterogeneity can be attributed to age, gender, experience with the product/service under study, physical characteristics, cognitive abilities, etc. leading to responses from different respondents being different. Thus, it would be ideal and practically more meaningful to incorporate the respondent effects in the model and then estimate the parameters of interest after eliminating the respondent effects.

A major concern with the traditional optimal choice designs is that the number of choice pairs to be shown to each respondent is very large even for moderately higher values of k and v_i. In a paired choice design, if the same set of N choice pairs are shown to the b respondents, then for the main effects to be estimable, N has to be at least $\sum_{i=1}^{k}(v_i - 1)$ $(= N_0$, say) and for the main plus two-factor interaction effects to be estimable, N has to be at least $\sum_{i=1}^{k}(v_i - 1) + \sum_{i=1}^{k-1}\sum_{j=i+1}^{k}(v_i - 1)(v_j - 1)$ $(= N_1$, say). Such optimal designs compromise on the quality of the responses. In the absence of practical optimal designs, [14] and [6] have obtained nearly optimal efficient designs having reasonably small N. Furthermore, the designs may also not remain optimal under a model which incorporates the respondent effects.

To overcome the difficulties as highlighted, we consider a model for the paired choice experiment incorporating the respondent effects. This approach enables the experimenter to use optimal designs after eliminating the respondent effects and provides an avenue to get optimal designs with reasonable number of choice pairs $s(< N)$ shown to b respondents. In what follows, considering respondents as blocks, s is called the block size of the paired choice design. Here, s can even be smaller than N_0 under the main effects model and can be smaller than N_1 under the main plus two-factor interaction effects model. Though under the traditional set-up one does not find practical optimal designs for estimating the main plus two-factor interaction effects, under our set-up we provide optimal designs with practical block sizes.

Without much rigor, instances of considering respondents as blocks have been reported in [1] and [11]. They highlight that 64% studies reported using a blocking column to allocate choice sets to respondents, 13% assigned choice sets randomly to respondents, 5% studies provided the full factorial to each respondent and for the remaining 18% of the studies, it couldn't be determined how choice sets were assigned to respondents. They also noted that the use of inappropriate blocking approaches jeopardizes the characteristics of the data.

Under the multinomial logit model or the linear paired comparison model, in Section 2 we first incorporate the respondent effects and then obtain the information matrix for estimating the parameters of interest eliminating the respondents effects. In Section 3, we provide optimal paired choice block designs for estimating the main effects for symmetric and asymmetric multi-level attributes with smaller number of choice pairs shown to each respondent. We also give a simple solution to the problem of identifying generators in the constructions of optimal paired choice designs. In Section 4, we provide optimal paired choice unblocked and block designs for estimating the main effects only under the main plus two-factor interaction effects model. Finally, in Section 5, we provide optimal paired choice block designs for estimating the main plus two-factor interaction effects.

2 Preliminaries and model incorporating respondent effects

Most of the work on optimal choice designs is based on the multinomial logit model approach of [12], or the multinomial logit model approach of [15]. The information matrices for estimating the parameters of interest have been obtained independently in both the approaches. [9] observed that the two approaches are equivalent for the purpose of finding optimal designs. For the sake of simplicity, we work with the multinomial logit model approach of [12]. The multinomial logit model supposes that the probability of preferring option 1 over option 2 in the ith choice pair can be expressed as $\pi_{12i} = e^{u_{1i}}/\{e^{u_{1i}} + e^{u_{2i}}\}$, where u_{1i} and u_{2i} represent the utilities attached to the two options in choice pair i. Similarly $\pi_{21i} = 1 - \pi_{12i}$ is the probability that option 2 is preferred over option 1. It follows that for each choice pair i the choice probabilities depend only on the utility difference $u_{1i} - u_{2i}$. For a design d with n choice pairs, since options are described by k attributes, the utilities are modelled using the linear predictor $U_j = P_{pj}\theta$, where θ is a $p \times 1$ vector representing the parameters of interest and P_{pj} is an effects-coded matrix for the jth option, $j = 1, 2$. The utility difference $U_1 - U_2 = (P_{p1} - P_{p2})\theta = P\theta$ is then a linear function of the parameter vector θ. Under the multinomial logit model of the equal choice probability or equivalently the indifference assumption of $\theta = 0$, the Fisher information matrix is $(1/4)P'P$.

Simultaneously, [8], [6] and [7] studied linear paired comparison designs which are analyzed under the linear paired comparison model. In the linear paired comparison model, rather than just choosing the preferred option, the respondents have to indicate on a scale how strong their preference is. The quantitative response Z is then the observed utility difference between the two options, which again depends on the difference matrix $P = P_{p1} - P_{p2}$. More precisely, the response is described by the model, $Z = U_1 - U_2 + \epsilon = (P_{p1} - P_{p2})\theta + \epsilon = P\theta + \epsilon$, where ϵ is the random error vector. The matrix $C = P'P$ is recognized as the information matrix under the linear paired comparison model. Since C is proportional to the information matrix under the indifference assumption of multinomial logit model, it follows that the designs optimal under the linear paired comparison model are also optimal under the multinomial logit model and vice versa. Optimal designs for estimating the main effects under the linear paired comparison model have been described in [6] and [7] and that under the equal choice probability multinomial logit model in [15], [4] and [2].

As noted in [9], most optimality results for choice designs and paired comparison designs are available for D-criterion since it is invariant to reparameterizations and is mathematically more tractable than other functions of the information matrix. Therefore, in what follows, by optimality we mean D-optimality. For a choice design T, let $0 < \gamma_1 \leq \gamma_2 \leq \cdots \leq \gamma_k$ be the eigenvalues of the information matrix. Then, a choice design d^* is said to be D-optimal if d^* maximizes the $\det(C)$.

For paired choice experiments, multinomial logit model as well as the linear paired comparison models are based on $U_1 - U_2$. Of the n choice pairs in a paired choice experiment, let each of the b respondents be shown s choice pairs. These s choice pairs are usually non repeated. By incorporating respondent effects, the relevant utility differences under the block model, with blocks being the respondents, becomes

$$U_1 - U_2 = (P_{p1} - P_{p2})\theta + W\beta = P\theta + W\beta, \tag{1}$$

where β represents the block effect, and $W = (w_{ij})$ is an $n \times b$ incidence matrix with $w_{ij} = 1$ if the ith choice pair belongs to the jth block and 0 otherwise. Without loss of generality, we take $W = I_b \otimes 1_s$, where I_p and 1_p denotes the identity matrix of order p and a $p \times 1$ vector of all ones, respectively.

In either multinomial logit model or linear paired comparison model, including respondent effects β can be regarded as adding b two level attributes to the set of p predictor variables. The corresponding difference matrix for the pairs in b blocks then has an additional component and can be written as (P, W). Then, under the indifference assumed multinomial logit block model, it follows that the information matrix for estimating θ and β is

$$M = \frac{1}{4}\begin{bmatrix} C & P'W \\ W'P & W'W \end{bmatrix} \tag{2}$$

where $C = P'P$ is as defined before. Moreover, upto a constant factor of $1/4$, M coincides with the information matrix in the extended linear paired comparison block model. Thus, optimal designs under the linear paired comparison block model are also optimal under the indifference assumed multinomial logit block model. The information matrix for estimating θ under the linear paired comparison block model after eliminating the block effect is

$$\tilde{C} = P'P - P'W(W'W)^{-1}W'P = C - (1/s)P'WW'P. \tag{3}$$

A choice block design is connected if all the parameters of interest are estimable, and this happens if and only if $\tilde{C}$ has rank p. In what follows, the class of all connected paired choice block designs with k attributes in b blocks each of size s is denoted by $\mathcal{D}_{k,b,s}$.

In order to study optimal paired choice designs for estimating the parameter θ under the block model, we first note that $P'WW'P$, and thus $C - \tilde{C}$ is a non-negative definite matrix. Therefore, from (3), if a choice design d is ϕ-optimal with respect to C (ϕ being a non-increasing optimality criterion), that is, under the unblocked model, then d is also ϕ-optimal with respect to $\tilde{C}$, that is, under the block model, provided $\tilde{C} = C$. An optimality criterion ϕ is non-increasing if $\phi(W_1) \leq \phi(W_2)$, whenever $W_1 - W_2$ is non-negative definite.

It is observed that eliminating respondent effects simultaneously controls for the within-pair order effects. Optimal designs under paired comparison model and the multinomial logit model in the presence of within-pair order effects have been obtained in [5] and [3] respectively.

3 Optimal designs under the main effects block model

Under the main effects block model, from (1) it follows that $U_1 - U_2 = (P_{M1} - P_{M2})\tau + W\beta = X\tau + W\beta$, where τ is a $\sum_{i=1}^{k}(v_i - 1) \times 1$ parameter vector for main effects, P_{Mj} is an $n \times \sum_{i=1}^{k}(v_i - 1)$ effects-coded matrix of the main effects for the jth option, $j = 1, 2$, and $X = P_{M1} - P_{M2}$. In a row of P_{Mj}, the levels of the ith attribute is represented by an effects-coded row vector of length $v_i - 1$. The level l is represented as a unit vector with 1 in $(l+1)$th position for $l = 0, \ldots, v_i - 2$, and level $v_i - 1$ is represented by -1 in each of the $v_i - 1$ positions, $i = 1, \ldots, k$. For example, for $v = 3$, effects-coded vectors for $l = 0, 1, 2$ are respectively $(1\ 0), (0\ 1)$ and $(-1\ \ -1)$.

From (3), the information matrix for estimating the main effects after eliminating the block effects is

$$\tilde{C}_M = C_M - (1/s)X'WW'X, \tag{4}$$

where $C_M = X'X$ is the information matrix for estimating the main effects. From (4), it follows that a necessary and sufficient condition for $\tilde{C}_M = C_M$ to hold is $W'X = 0$. Therefore, by suitably allocating the optimal choice pairs in blocks such that $W'X = 0$, an optimal paired choice block design is obtained. Let $P_{Mj} = ((P_j)'_1 \cdots (P_j)'_t \cdots (P_j)'_b)'$ where $(P_j)_t$ represents P_{Mj} for the tth block. Then the condition $W'X = 0$ is equivalent to the condition $1'(P_1)_t = 1'(P_2)_t$, $t = 1, \ldots, b$.

Theorem 3.1. *A necessary and sufficient condition for $\tilde{C}_M = C_M$ to hold is that for each block, the frequency distribution of the levels of every attribute involved in the two options of choice pairs are the same for each of the two options.*

Proof. Let $(P_j)_t = ((P_j)_t^1 \cdots (P_j)_t^w \cdots (P_j)_t^k)$ where $(P_j)_t^w$ is of order $s \times (v_w - 1)$ and represents $(P_j)_t$ for the wth attribute. Therefore for $t = 1, \ldots, b$, if $1'(P_1)_t = 1'(P_2)_t$, then $1'(P_1)_t^w = 1'(P_2)_t^w$ for every w and t.

Now, since the ith column of $(P_j)_t^w$ provides frequency of level i and level v_w in the wth attribute of the jth option in the tth block, therefore, $1'(P_1)_t^w = 1'(P_2)_t^w$ implies that the frequency of each of the levels of attribute w are same in the two options among the s choice pairs in block t.

The sufficiency condition follows trivially since for each block, the frequency of the levels of every attribute being same for each of the two options in s choice pairs implies that $1'(P_1)_t = 1'(P_2)_t$ for every t. ∎

We have the following two corollaries.

Corollary 3.1. *$\tilde{C}_M = C_M$ if the levels of every attribute appear equally often in both the options in every block.*

Corollary 3.2. *When $s = 2$, for any X such that $W'X = 0$ implies that the paired choice block design has repeated choice pairs.*

Proof. From Theorem 3.1, $W'X = 0$ implies that in a block, for every attribute, each level should occur same number of times in first and second option. Since $s = 2$, to satisfy Theorem 3.1, if an attribute in option 1 has $(v_1, v_2)'$, same attribute in option 2 would have $(v_2, v_1)'$ leading to corresponding pairs of choice pairs as (v_1, v_2) and (v_2, v_1). Since this is true for all the attributes, one would have repeated choice pairs in a block. ∎

Under the linear paired comparison model, a design d optimally estimates the main effects if the information matrix $(1/4)C_M$ is a block diagonal matrix having k blocks such that $C_M = diag(C_{(1)}, \ldots, C_{(k)})$ where $C_{(i)} = z_i(I_{v_i-1} + J_{v_i-1})$ with $z_i = 2n/(v_i - 1)$, $i = 1, \ldots, k$.

Let H_m denote a Hadamard matrix of order $m \equiv 0 \pmod 4$. An orthogonal array $OA(n, k, v_1 \times \cdots \times v_k, t)$, having n rows, $k(\geq 2)$ columns, $v_1, \ldots, v_k$ symbols and strength $t(\leq n)$, is an $n \times k$ array, with elements in the ith column from a set of v_i distinct symbols $(i = 1, \ldots, k)$, in which all possible combinations of symbols appear equally often as rows in every $n \times t$ subarray. An orthogonal array is symmetric if $v_i = v$ for all i and the corresponding OA is denoted by $OA(n, k, v^k, t)$, else it is an asymmetric orthogonal array. The link `http://support.sas.com/techsup/technote/ts723.html` and [10] provides a comprehensive summary of orthogonal arrays and their constructions.

[15], [4] and [2] provide the $OA + G$ method for constructing optimal paired choice designs using orthogonal arrays of strength 2 and generators G. From an $OA(n, k, v_1 \times \cdots \times v_k, 2)$, an optimal paired choice design is constructed using $h = lcm(h_1, \ldots, h_k)$

generators where h_i is the number of generators corresponding to attribute i having v_i levels and $lcm(a_1, \ldots, a_k)$ denotes the least common multiple of $a_1, \ldots, a_k$. Note that corresponding to the attribute level v_i, the h_i generators are $g_j = (0, j), j = 1, \ldots, h_i$.

Since $OA + G$ method entails adding generators to the orthogonal array of strength 2, the off-diagonal elements of $X'X$ corresponding to two different attributes is zero since under each level of the first attribute, all the levels of the second attribute occur equally often. Also, since under each column (attribute), levels are equally replicated in an orthogonal array, to see that each $C_{(i)}, i = 1, \ldots, k$ attains optimal structure of the form $z_i(I_{v_i-1} + J_{v_i-1})$, it is enough to show that $X'X$ corresponding to a paired choice design with one factor, say at v levels, attains the structure $z(I_{v-1} + J_{v-1})$ where $z = 2n/(v-1)$. In fact, since in an orthogonal array, every level occurs equally often, it is enough to show the optimal structure of $X'X$ to be $z(I_{v-1} + J_{v-1})$ where $n = v$.

In the literature, arriving at the generators G have been usually through a trial and error approach and no systematic results on the structure of such generators, in general, appear to exist. In fact, [2] highlight the complexities involved in choosing the sets of generators. We provide a simple result which gives the generators for any v.

Theorem 3.2. *To get an optimal paired choice design through the $OA+G$ method, the number and structure of generators for a typical attribute at v levels is $h = v - 1, g_j = (0, j), j = 1, \ldots, v - 1$ for v even and $h = (v - 1)/2, g_j = (0, j), j = 1, \ldots, (v - 1)/2$ for v odd.*

Proof. While using the generator g_j, let P_1^0, P_2^j be the $v \times 1$ effects-coded matrix for the main effects for first and second option respectively, corresponding to any one attribute at v levels. When $h > 1$, note that X is the collection of different matrices generated out of the corresponding $\{P_1^0, P_2^j\}, j = 1 \ldots, h$ of choice pairs. For simplicity, we denote P_1^0 by P_0 and P_2^j by P_j, $j = 1, \ldots, v - 1$. Also, note that $1'P_j = 0$ and $\sum_{j=0}^{v-1} P_j = 0$.

Consider the information matrix $X'X$ for v even. $X'X = \sum_{j=1}^{v-1}(P_0 - P_j)'(P_0 - P_j) = \sum_{j=1}^{v-1}(P_0'P_0 + P_j'P_j - P_0'P_j - P_j'P_0) = \sum_{j=1}^{v-1}\{2(I_{v-1} + J_{v-1})\} - P_0'(\sum_{j=1}^{v-1} P_j) - (\sum_{j=1}^{v-1} P_j')P_0 = \{2(v-1)(I_{v-1} + J_{v-1})\} - P_0'(-P_0) - (-P_0')P_0 = 2\{(v-1)(I_{v-1} + J_{v-1})\} + 2P_0'P_0 = 2v(I_{v-1} + J_{v-1})$. Thus, for v even, $h = v - 1$ generators of the type $g_j = (0, j), j = 1, \ldots, v - 1$ leads to optimal structure of $X'X$.

For v odd, $P_j'P_0 = P_0'P_{v-j}$. To see this, we note that, if say, mth row of P_0 corresponds to the level i, then the mth row of P_{v-j} corresponds to the level $i-j$ (mod v) and similarly, if say, lth row of P_j corresponds to the level i, then the lth row of P_0 corresponds to the level $i - j$ (mod v). This makes the lth row of P_j and P_0 same as the mth row of P_0 and P_{v-j} for every two rows $l \neq m = 1, \ldots, v$. Therefore, $X'X = \sum_{j=1}^{(v-1)/2}(P_0 - P_j)'(P_0 - P_j) = \sum_{j=1}^{(v-1)/2}(P_0'P_0 + P_j'P_j - P_0'P_j - P_j'P_0) = \sum_{j=1}^{(v-1)/2}\{2(I_{v-1} + J_{v-1})\} - \sum_{j=1}^{(v-1)/2}(P_0'P_j + P_j'P_0) = (v-1)(I_{v-1} + J_{v-1}) - \sum_{j=1}^{(v-1)/2}(P_0'P_j + P_0'P_{v-j}) = (v-1)(I_{v-1} + J_{v-1}) - P_0'\sum_{j=1}^{(v-1)/2}(P_j + P_{v-j}) = (v - 1)(I_{v-1} + J_{v-1}) - P_0'\sum_{j=1}^{v-1} P_j = (v - 1)(I_{v-1} + J_{v-1}) - P_0'(-P_0) = v(I_{v-1} + J_{v-1})$. Thus, for v odd, $h = (v - 1)/2$ generators of the type $g_j = (0, j), j = 1, \ldots, (v - 1)/2$ leads to optimal structure of $X'X$. ■

For example, for $v = 5$, we have $h = 2$, $g_j = (0, j), j = 1, 2$ and for $v = 6$, $h = 5$, $g_j = (0, j), j = 1, 2, 3, 4, 5$. Suppose there are 3 attributes with $v_1 = 2, v_2 = 3$ and $v_3 = 4$. Then $h_1 = 1, h_2 = 1, h_3 = 3$ and $h = 3$. This leads to generators being $G_1 = (000, 111), G_2 = (000, 112)$ and $G_3 = (000, 113)$. Thus, for a given $OA(n_1, 3, 2 \times 3 \times 4, 2)$, the corresponding optimal paired choice design with parameters k, $v_1 = 2, v_2 = 3, v_3 = 4$, $b = 1$, $s = 3n_1$, is obtained using the $OA + G$ method with 3 generators. In general, for a given $OA(n_1, k, v_1 \times \cdots \times v_k, 2)$, the corresponding optimal paired choice design d_1 with parameters k, $v_1, \ldots, v_k$ $b = 1$, $s = hn_1$, is obtained using the $OA + G$ method with generators $G_i, i = 1, \ldots, h$.

Theorem 3.3. *If an $OA(n_1, k+1, v_1 \times \cdots \times v_k \times \delta, 2)$ exists, then the corresponding optimal paired choice design d_1 (with parameters k, $v_1, \ldots, v_k$, $b = 1$, $s = hn_1$) using the method of $OA + G$ with h generators exists and a resultant paired choice block design d_2 with parameters k, $v_1, \ldots, v_k$, $b = h\delta$, $s = n_1/\delta$ is optimal in $\mathcal{D}_{k,b,s}$.*

Proof. For a given $OA(n_1, k+1, v_1 \times \cdots \times v_k \times \delta, 2)$, corresponding to the k attributes at levels $v_i, i = 1, \ldots, k$ and $h = lcm(v_1, \ldots, v_k)$, there exists an optimal paired choice design d_1 (with parameters k, $v_i, i = 1, \ldots, k$, $b = 1$, $s = hn_1$). From d_1, the choice pairs obtained through each of the h generators constitute a block of size n_1, since n_1 rows of a block form the orthogonal array in the first option and, with labels re-coded through the generator, in the second option. Now, to further reduce the block size, using the attribute at level δ as a blocking attribute, we observe that for every attribute in each of the blocks with $s = n_1/\delta$, each level occurs equally often. Therefore, by Corollary 3.1, d_2 is optimal in the class of all paired choice block designs with parameters k, $v_1, \ldots, v_k$, $b = h\delta$, $s = n_1/\delta$. ■

The following corollary is an immediate consequence of Theorem 3.3 by considering $\delta = 1$ in $OA(n_1, k+1, v_1 \times \cdots \times v_k \times \delta, 2)$.

Corollary 3.3. *If an $OA(n_1, k, v_1 \times, \ldots, \times v_k, 2)$ exists, then a resultant paired choice block design d_3 with parameters k, $v_1, \ldots, v_k$, $b = h$, $s = n_1$ is optimal in $\mathcal{D}_{k,b,s}$.*

Example 3.1. *From $OA(24, 15, 2^{13} \times 3 \times 4, 2)$, an optimal paired choice design d_1 for estimating $k = 15$ main effects with ith attribute at 2 levels, $i = 1, \ldots, 13$, 14th attribute at 3 level and 15th attribute at 4 level is obtained by using $h = 3$ generators ($h_i = 1, i = 1, \ldots, 14$, and $h_{15} = 3$). Note that the generators are $G_1 = (000, 111), G_2 = (000, 112), G_3 = (000, 113)$*

If an optimal paired choice block design is needed for estimating $k = 14$ main effects, the following three situations may arise.

(i) 13 attributes at 2 levels and 1 attribute at 3 level: For $k = 14$, $v_1 = \cdots = v_{13} = 2, v_{14} = 3$, there are three designs: d_1 with $b = 1, s = 24$, d_2 with $b = 4, s = 6$, and d_3 with $b = 1, s = 24$, .

(ii) 13 attributes at 2 levels and 1 attribute at 4 level: For $k = 14$, $v_1 = \cdots = v_{13} = 2, v_{14} = 4$, there are three designs: d_1 with $b = 1, s = 72$, d_2 with $b = 9, s = 8$ and d_3 with $b = 3, s = 24$.

(iii) 12 attributes at 2 levels, 1 attribute at 3 level and 1 attribute at 4 level: For $k = 14$, $v_1 = \cdots = v_{12} = 2, v_{13} = 3, v_{14} = 4$, *there are three designs:* d_1 *with* $b = 1, s = 72$, d_2 *with* $b = 6, s = 12$ *and* d_3 *with* $b = 3, s = 24$.

For estimating all 15 attributes, d_3 *with parameters* $k = 15$, $v_1 = \cdots = v_{13} = 2, v_{14} = 3, v_{15} = 4, b = 3, s = 24$ *is optimal.*

As an illustration, consider $2^4 \times 3$ *paired choice block design with parameters* $k = 5, b = 4, s = 6$.

$d =$

B_1		B_2		B_3		B_4	
(00000)	(11111)	(01102)	(10010)	(10112)	(01000)	(10001)	(01112)
(11010)	(00101)	(11110)	(00001)	(00111)	(11002)	(00012)	(11100)
(01101)	(10012)	(11011)	(00102)	(01002)	(10110)	(10100)	(01011)
(11002)	(00110)	(00100)	(11011)	(11101)	(00012)	(01011)	(10102)
(10111)	(01002)	(10012)	(01100)	(01010)	(10101)	(01110)	(10001)
(00112)	(11000)	(00001)	(11112)	(10000)	(01111)	(11102)	(00010)

For many parameter sets corresponding to k attributes each at v levels, [7] and [4] provided constructions of optimal paired choice designs with reduced number of choice pairs as against the $OA+G$ method. Note that only Construction 3.2 and Construction 3.4 of [4] are applicable for paired choice designs and that Construction 3.4 reduces to the method in Section 5 of [7] for paired choice designs. We now show how an optimal paired choice block design can be constructed from these two construction methods.

Theorem 3 of [7] states that from $m \geq k$ rows of H_m, an optimal paired choice design d_4 with parameters $k, v, b = 1, s = mv(v-1)/2$ is constructed using the $v(v-1)/2$ combinations of v symbols taken two at a time. If v is odd, $(v-1)/2$ is an integer and the $v(v-1)/2$ combinations can be arranged in rows such that each of the two columns have every symbol appearing equally often. This is always possible and follows using systems of distinct representatives. Now, for every row of $\{H_m, -H_m\}$, $v(v-1)/2$ choice pairs are obtained where a choice pair is constructed by replacing '1' in $\{H_m, -H_m\}$ by first value of the combination and '-1' in $\{H_m, -H_m\}$ by the corresponding second value of the combination. Therefore, using $v(v-1)/2$ choice pairs, corresponding to each of the rows of $\{H_m, -H_m\}$, as a block, a paired choice block design d_5 with parameters $k, v, b = m, s = v(v-1)/2$ is obtained which following Corollary 3.1 is optimal.

Theorem 3.4. *If a Hadamard matrix* H_m *exists, then an optimal paired choice design* d_4 *with parameters* $k, v, b = 1, s = mv(v-1)/2$, $k \leq m$ *exists. Furthermore for* v *odd, a paired choice block design* d_5 *with parameters* $k, v, b = m, s = v(v-1)/2$ *exists and is optimal in* $\mathcal{D}_{k,b,s}$.

Example 3.2. *Consider* $v = 3$ *with combinations* $(0,1), (1,2), (2,0)$ *and the Hadamard matrix* H_4. *An optimal paired choice design* d_4 *with parameters* $k = 4, v = 3, b = 1, s = 12$ *exists. Furthermore, since* v *is odd, an optimal paired choice block design* d_5 *is constructed with parameters* $k = 4, v = 3, b = 4, s = 3$ *by replacing 1s and* -1*s respectively by* $(0,1), (1,2), (2,0)$ *in each of the four rows of* $\{H_4, -H_4\}$ *and by considering choice pairs generated by each row of* $\{H_4, -H_4\}$ *as a block.*

$d_5 =$

B_1		B_2		B_3		B_4	
(0000)	(1111)	(0101)	(1010)	(0011)	(1100)	(0110)	(1001)
(1111)	(2222)	(1212)	(2121)	(1122)	(2211)	(1221)	(2112)
(2222)	(0000)	(2020)	(0202)	(2200)	(0022)	(2002)	(0220)

Similarly, Construction 3.2 of [4] uses an $OA(n_2, k+1, v^k \times v_{k+1}, 2)$ with $v_{k+1} = n_2/v$ and forms v_{k+1} parallel sets with v rows each. Then, an optimal paired choice design d_6 with parameters $k, v, b = 1, s = v_{k+1}\binom{v}{2}$ is constructed using the $v(v-1)/2$ combinations of v symbols taken two at a time. Again as earlier, for v odd, the $v(v-1)/2$ combinations can be arranged in rows such that each of the two columns have every symbol appearing equally often. For each such combination, corresponding rows from each of the v_{k+1} parallel sets are chosen to form the choice pairs of the optimal paired choice design d_6. Considering each parallel class with $v(v-1)/2$ choice pairs as a block, a paired choice block design d_7 with parameters $k, v, b = v_{k+1}, s = v(v-1)/2$ is optimal using Corollary 3.1.

Theorem 3.5. *If an $OA(n_2, k+1, v^k \times v_{k+1}, 2)$ with $v_{k+1} = n_2/v$ exists, then an optimal paired choice design d_6 with parameters $k, v, b = 1, s = n_2(v-1)/2$ exists. Furthermore for v odd, a paired choice block design d_7 with parameters $k, v, k, b = n_2/v, s = v(v-1)/2$ exists and is optimal in $\mathcal{D}_{k,b,s}$.*

Example 3.3. *From an $OA(9, 4, 3^4, 2)$, an optimal paired choice design d_6 with parameters $k = 3, v = 3, b = 1, s = 9$ is constructed. Furthermore, since v is odd, constructing 3 parallel sets using any column as a blocking attribute and by taking $(1,2), (2,3), (3,1)$ rows of orthogonal arrays as choice pairs, 3 choice pairs are formed for each parallel sets. As a result, an optimal paired choice block design d_7 with parameters $k = 3, v = 3, b = 3, s = 3$ is obtained.*

$d_7 =$

B_1		B_2		B_3	
(000)	(121)	(022)	(110)	(011)	(102)
(121)	(212)	(110)	(201)	(102)	(220)
(212)	(000)	(201)	(022)	(220)	(011)

Note that using $OA(9, 4, 3^4, 2)$, the $OA+G$ method allows one to get an alternate optimal paired choice block design with the same parameters $k = 3, v = 3, b = 3, s = 3$.

Remark 3.1. *For v even, Theorem 3.4 and Theorem 3.5 do not give optimal paired choice block designs. However, as a substitute, one can always construct a paired choice block design from the $OA+G$ construction method using Theorem 3.3 and Corollary 3.3. For instance, for $v = 4, 10 \leq k \leq 12$, as in [4], an optimal paired choice design d_1, d_4 and d_6 can be constructed in 144, 72 and 72 choice pairs respectively. Since $v = 4$ is even, we can only construct an optimal paired choice block design using $OA+G$ method with starting orthogonal array $OA(48, 13, 4^{12} \times 12^1, 2)$. Since for $v = 4$, the number of generators $h = 3$, using the attribute at 12 levels as the blocking attribute, from Theorem 3.3, d_2 with $k = 12, v = 4, b = 36, s = 4$ is optimal in $\mathcal{D}_{k,b,s}$. Also, using Corollary 3.3, a paired choice block design d_3 with $k = 12, v = 4, b = 3, s = 48$ is also optimal in $\mathcal{D}_{k,b,s}$.*

For any $t (\geq 1)$, we let $b = t_i = it$. Here, the number of blocks b is t repetitions of i blocks of size s. Thus, the number of respondents in a choice experiment is a multiple of i.

Table 1: Feasible number of pairs s per respondent

v	k	Traditional	Proposed
2	3	4	4
2	4	4*	4
2	5,6	8	4,6,8
2	7	8	4,6,8
2	8	8*	4,6,8
2	9,10	12	4,6,8,10,12
2	11	12	4,6,8,10,12
2	12	12*	4,6,8,10,12
3	3	9,12	3,6,9,12
3	4	9,12,18	3,6,9,12,18
3	5,6	18,24	3,6,9,12,18,24
3	7	18,24,27	3,6,9,12,18,24,27
3	8	24,27	3,6,9,12,18,24,27
3	9	27,36	3,6,9,12,18, 27, 36
3	10-12	27,36	3,6,9,12,18, 27, 36
4	3-4	24*,48	4,8,12,16,24,48
4	5	48	4,8,12,16,24, 32,48, 96
4	6-8	48*,96	4,8,12,16,24, 32,48, 96
4	9	72*,96	4,8,12,16,24,32,36,48,72,96
4	10-12	72*,144	4,8,12,16,24,32,36,48,72,96
5	3-4	40,50	5,10,20,25,40,50
5	5	50,80	5,10,20,25,40,50,80
5	6	50,80,100	5,10,20,25,40,50,80,100
5	7-8	80,100	5,10,20,25,40,50,80,100
5	9-10	100,120	5,10,20,25,30,40,50,60,100,120
6	3	60*,180	12,18, 36, 60,180
6	4	60*,180*,360	6,12,18,24,30,36,60,72,90,120,180,360
6	5-6	120*,180*,360	6,12,18,24,30,36,60,72,90,120,180,360
7	3-4	84,147	7,21,42,49,84,147
7	5-7	147,168	7,21,42,49,84,147
7	8	147,168,294	7,14,21,42,49,84,98,147,168,294

[13] provided optimal paired choice designs for $v = 2$ when $i = 1, t = 1$ and $s \not\equiv 0 \pmod 4$. However, when $t > 1$ (which is the more practical set-up since the blocks or respondents are many in a choice experiment), the designs are no longer optimal. The designs provided through Theorem 3.3 generate optimal paired choice designs for $v = 2$ when $s \equiv 2 \pmod 4$ and $i = 2$.

We present in Table 1, the number of choice pairs per respondent of the traditional designs as listed in Table 2 of [4] and the corresponding block sizes obtainable under our constructions based on Theorem 3.3, Theorem 3.4 and Theorem 3.5. In Table 1, the traditional designs marked $*$ are not optimal under the block set-up for blocks of size s and $b = t_1$. However, by having number of blocks t_i with $i > 1$, optimal designs having blocks of size s is obtainable as presented under the Proposed column. In Table 2, we list the values of s and i corresponding to optimal designs obtained through Theorem 3.3 and Theorem 3.4. It is observed that in the parameter range of Table 2, Theorem 3.5 does not provide any additional designs not obtainable from other Theorems in this section.

Table 2: Optimal designs in $\mathcal{D}_{k,t_i,s}$

v	k	Theorem 3.3 (s,i)	Theorem 3.4 (s,i)
2	3	(4,1)	
2	4	(4x,2/x), x=1,2	
2	5-6	(4x,2/x), x=1,2 (6x,2/x), x=1,2	
2	7	(8,1) (6x,2/x), x=1,2 (4x,4/x), x=1,2,4	
2	8	(6x,2/x), x=1,2 (4x,4/x), x=1,2,4	
2	9-10	(6x,2/x), x=1,2 (4x,4/x), x=1,2,4 (10x,2/x), x=1,2	
2	11	(12,1) (10x,2/x), x=1,2 (4x,4/x), x=1,2,4 (6x,4/x), x=1,2,4	
2	12	(10x,2/x), x=1,2 (4x,4/x), x=1,2,4 (6x,4/x), x=1,2,4	
3	3	(3x,3/x), x=1,3	(3x,4/x), x=1,2,4
3	4	(9,1) (3x,6/x), x=1,2,3,6	(3x,4/x), x=1,2,4
3	5,6	(3x,6/x), x=1,2,3,6	(3x,8/x), x=1,2,4,8
3	7	(9x,2/x), x=1,2 (3x,9/x), x=1,3,9	(3x,8/x), x=1,2,4,8
3	8	(3x,9/x), x=1,3,9	(3x,8/x), x=1,2,4,8
3	9	(3x,9/x), x=1,3,9	(3x,12/x), x=1,2,3,4,6,12
3	10-12	(9x,3/x), x=1,3 (3x,12/x), x=1,2,3,4,6,12	(3x,12/x), x=1,2,3,4,6,12
4	3-4	(4x,12/x), x=1,2,3,4,6,12	
4	5	(16x,3/x), x=1,3 (4x,24/x), x=1,2,3,4,6,8,12,24	
4	6-8	(4x,24/x), x=1,2,3,4,6,8,12,24	
4	9	(16x,6/x), x=1,2,3,6 (4x,36/x), x=1,2,3,4,6,9,12,18,36	
4	10-12	(4x,36/x), x=1,2,3,4,6,9,12,18,36	
5	3-4	(5x,10/x), x=1,2,5,10	(10x,4/x), x=1,2,4
5	5	(5x,10/x), x=1,2,5,10	(10x,8/x), x=1,2,4,8
5	6	(25x,2/x), x=1,2 (5x,20/x), x=1,2,4,5,10,20	(10x,8/x), x=1,2,4,8
5	7-8	(5x,20/x), x=1,2,4,5,10,20	(10x,8/x), x=1,2,4,8
5	9-10	(5x,20/x), x=1,2,4,5,10,20	(10x,12/x), x=1,2,3,4,6,12
6	3	(12x,15/x), x=1,3,5,15 (18x,10/x), x=1,2,5,10	
6	4	(6x,60/x), x=1,2,3,4,5,6,10,12,15,20,30,60	
6	5-6	(6x,60/x), x=1,2,3,4,5,6,10,12,15,20,30,60	
7	3-4	(7x,21/x), x=1,3,7,21	(21x,4/x), x=1,2,4
7	5-7	(7x,21/x), x=1,3,7,21	(21x,8/x), x=1,2,4,8
7	8	(49x,3/x), x=1,3 (7x,42/x), x=1,2,3,6,7,14,21,42	(21x,8/x), x=1,2,4,8

4 Optimal designs under the broader main effects block model

In this section, we consider estimation of main effects under the broader main effects model for a v^k choice design. Broader main effects model includes the main effects and the two-factor interaction effects with our interest only in the estimation of the main effects. Optimal designs under the broader main effects model have been recently obtained in [13].

From (1), the relevant utility differences with the introduction of respondent effects become

$$U_1 - U_2 = (P_{M1} - P_{M2})\tau + (P_{I1} - P_{I2})\gamma + W'\beta = X\tau + Y\gamma + W'\beta, \tag{5}$$

where γ is a $k(k-1)(v-1)^2/2 \times 1$ parameter vector for two-factor interaction effects, P_{Ij} is an $n \times k(k-1)(v-1)^2/2$ effects-coded matrix of the two-factor interaction effects for the jth option, $j = 1, 2$, X, τ, W and β are as defined in Section 3. Let $P_{Ij} = (P_{Ij}^{1'}, \ldots, P_{Ij}^{n'})'$ where P_{Ij}^i corresponds to P_{Ij} for the ith choice pair. Representing the columns of P_{Mj} corresponding to the ith choice pair and αth attribute by $P_{Mj(\alpha)}^i$, we get $P_{Ij}^i = (P_{Mj(1)}^i \otimes P_{Mj(2)}^i, P_{Mj(1)}^i \otimes P_{Mj(3)}^i, \ldots, P_{Mj(k-1)}^i \otimes P_{Mj(k)}^i)$.

From (2),the information matrix for estimating the main effects after eliminating for the two-factor interaction effects and the block effects is

$$\tilde{C}_B = C_M - [X'Y \; X'W] \begin{bmatrix} Y'Y & Y'W \\ Y'W & W'W \end{bmatrix}^- \begin{bmatrix} Y'X \\ W'X \end{bmatrix}.$$

Therefore, for a design which is optimal under the unblocked model is optimal under the broader main effects block model if $\tilde{C}_B = C_M$, i.e., both $Y'X = 0$ and $W'X = 0$. Note that the optimal designs in Section 3 already satisfy $W'X = 0$ under the 2^k choice design set-up. [13] showed that for 2^k choice experiment, a choice design satisfies $Y'X = 0$ provided the effects-coded matrix X has elements ± 1 only. Thus, the designs obtained for the 2^k choice design set-up in Theorem 3.3 are also optimal for paired choice designs under the broader main effects block model since X has elements ± 1 only.

For $v \geq 3$, there do not exist any corresponding result for v^k paired choice designs under the broader main effects model. As such, for $v \geq 3$, the designs under the main effects model, as provided in Section 3, do not remain optimal under the broader main effects model. Thus, we first provide optimal paired choice designs for v^k choice designs under the broader main effects model for $v \geq 3$. Thereafter, we find the optimal designs under the broader main effects block model.

Theorem 4.1. *If an $OA(n_1, k, v^k, 3)$ exists, then there exists an optimal paired choice design d_1^B with parameters $k, v, s = hn_1, b = 1$.*

Proof. A paired choice design using $OA + G$ method is optimal under the broader main effects model if a design is obtained using the generators as in Section 3, since that leads to an optimal C_M, and the design has $X'Y = 0$.

Since strength 3 orthogonal array is a strength two orthogonal array as well, it automatically has optimal C_M. Additionally, for $X'Y = 0$, we verify below the two equivalent conditions that hold.

(i) *For h attribute in X, the inner product of a hth column of X and a corresponding column of Y representing two-factor interaction effects of hth attribute with any other attribute is zero:* This is true since any strength 3 orthogonal array is a strength two orthogonal array and thus for every distinct value in hth column, each of the v levels are coming equally often in the other attribute. Thus for every value of h, the corresponding column sums of Y are 0, which in turn leads to corresponding column sum of $X'Y$ to be 0.

(ii) *For h attribute in X, the inner product of a hth column of X and a corresponding column of Y representing two-factor interaction effects between l_1th and l_2th attribute is zero:* Since starting array is a strength 3 orthogonal array, for every value of h, l_1 and l_2 will take each of the different v^2 values. It is easy to see that for each of the v^2 values, corresponding column sums of P_{I1} and P_{I2} and hence Y are zero. Since this hold for every value in hth attribute, it leads to corresponding column sum of $X'Y$ to be 0. ■

As before in Section 3, if a paired choice design has h generators, then an optimal paired choice block design d_2^B can be constructed with parameters $k, v, s = n_1, b = h$ considering pairs generated by different generators as blocks.

Theorem 4.2. *If an $OA(n_1, k, v^k, 3)$ exists, then a paired choice block design d_2^B with parameters $k, v, s = n_1, b = h$ constructed from d_1^B, is optimal in $\mathcal{D}_{k,b,s}$.*

If $n_1 = v^k$, then an $OA(v^k, k, v^k, k)$ exists by taking a complete factorial design. If $k \geq 3$, then the orthogonal array or the complete factorial design is a strength 3 orthogonal array and thus satisfies the conditions in Theorem 4.1 and Theorem 4.2. A paired choice block design $d_3^B \in \mathcal{D}_{k,b,s}$ with parameters $k, v, s = v^2, b = hv^{k-2}$ is constructed as follows:

(i) Keep any one column fixed at any level say x, and then group together any set of v rows in one block.

(ii) Then $v - 1$ sets each of v rows are obtained from this initial block cyclically by adding 1 (mod v),..., $(v - 1)$ (mod v).

(iii) This forms a block.

(iv) Repeat the procedure for each set of v non-repeated rows for level x of fixed column to get v^{k-2} blocks each of size v^2.

(v) Note that paired choice design so obtained has non repeated choice pairs in each block.

Theorem 4.3. *Starting from $OA(v^k, k, v^k, k)$, there exists an optimal paired choice block design d_2^B with parameters $k, v, s = v^k, b = h$. Furthermore, the paired choice block design $d_3^B \in \mathcal{D}_{k,b,s}$ with parameters $k, v, s = v^2, b = hv^{k-2}$ is optimal in $\mathcal{D}_{k,b,s}$.*

If $n_1 < v^{k-1}$, and an orthogonal array $OA(n_1, k, v^k, 3)$ exists, then we use this orthogonal array to optimize on the number of choice pairs. We now show that, since d_2^B exists for the $OA(n_1, k, v^k, 3)$, there exists an optimal paired choice block design $d_4^B \in \mathcal{D}_{k,b,s}$ with parameters $k, v, s = v^2, b = hn_1/v$.

(i) Generate $v-1$ copies by adding 1 (mod v),..., $(v-1)$ (mod v) to the base design d_2^B.

(ii) Since same element is added to each of the entries of d_2^B, it is just the recoding of the same design and hence, the resultant design with $n = n_1 v$ is also optimal under paired choice block model.

(iii) Furthermore, a block is obtained by picking $\{i_1, \ldots, i_v\}$th, $i_x = 1, \ldots, n_1$ rows from each of the v copies of the orthogonal array.

(iv) Repeat the procedure for each of the v-sized sets of n_1 initial rows. Note that once a row is attached to one block, it cannot be considered as part of another block.

(v) We also note that the blocks obtained in this way have each level occurring equally often under each attributes and the blocks and hence the design d_4^B is optimal in $\mathcal{D}_{k,b,s}$ using Corollary 3.1.

Therefore, we have the following result.

Theorem 4.4. *If a $OA(n_1, k, v^k, 3)$ and the corresponding d_2^B exists, then there exists a paired choice block design $d_4^B \in \mathcal{D}_{k,b,s}$ with parameters $k, v, s = v^2, b = hn_1/v$ which is optimal in $\mathcal{D}_{k,b,s}$.*

5 Optimal block designs for estimating the main effects and the two-factor interaction effects

The literature on optimal paired choice designs for estimating the main plus two-factor interaction effects is very limited since such designs require a large number of choice pairs to be shown to every respondent. However, [6] and [15] have provided optimal designs under this set-up for k attributes each at two levels. Let $q = \lceil k/2 \rceil$, where $\lceil z \rceil$ represents the smallest integer greater than or equal to z. The construction method typically entails starting with an orthogonal array $OA(n_2, k, 2^k, 4)$ and then taking the foldover of α attributes in second option, keeping the rest of the $k - \alpha$ attributes same where $\alpha = q$ for k odd and $\alpha = q$ and $q+1$ for k even. Here, foldover of an attribute in the second option of a choice pair means that the attribute level in second option is different from that of the first option. Thus, for each of the n_2 rows of the orthogonal array one would have $\binom{k}{\alpha}$ choice pairs, and hence, such a paired choice design d_1^I with parameters $k, v, s, b = 1$ is optimal where $s = n_2\binom{k}{q}$ for k odd and $s = n_2\binom{k+1}{q+1}$ for k even. The information matrix for an optimal paired choice design d_1^I is $2(k+1)/kI_{k+k(k-1)/2}$ for k odd and $2(k+2)/(k+1)I_{k+k(k-1)/2}$ for k even.

Incorporating respondent effects, the model is as given in (5). However, in contrast to Section 4, interest here lies in the estimation of both the main-effects and the two-factor interaction effects. The information matrix for estimating the main plus two-factor interaction effects under the multinomial logit model incorporating respondent effects is

$$\tilde{C} = \begin{bmatrix} X'X & X'Y \\ Y'X & Y'Y \end{bmatrix} - (1/s)[X'W \; Y'W] \begin{bmatrix} W'X \\ W'Y \end{bmatrix}.$$

As earlier, to achieve optimal block designs, we start with an optimal paired choice design d_1^I and enforce blocking such that $W'X = 0$ and $W'Y = 0$.

Let pair (a, b), means that a and b are the levels corresponding to the attributes for the first and second options respectively.

Theorem 5.1. *An optimal paired choice design d_1^I can be blocked to get an optimal paired choice block design if and only if for every block,*

(i) Under an attribute, the frequency of the pair $(1, 0)$ is same as the frequency of the pair $(0, 1)$;

(ii) Under any two attributes, the number of elements in a pair from the set $\{(01, 00),$ $(01, 11),$ $(10, 00), (10, 11)\}$ is same as the number of elements from set $\{(00, 01), (00, 10),$ $(11, 01), (11, 10)\}$.

Proof. From Theorem 3.1, $W'X = 0$ in d_1^I if and only if for every attribute having foldover in the second option of a choice pair, levels 0 and 1 appear equally often in both the options in every block and thus, the frequency of the pair $(1, 0)$ is same as the frequency of the pair $(0, 1)$ under every attribute in each block.

Let $Y = (Y_1' \cdots Y_t' \cdots Y_b')'$ with $Y_t = (P_{I1})_t - (P_{I2})_t$ being the $s \times k$ matrix corresponding to tth block. Then, the condition $W'Y = 0$ is equivalent to the condition $1'(P_{I1})_t = 1'(P_{I2})_t$ for every $t = 1, \ldots, b$ where $(P_j)_t$ represents P_{Ij} for the tth block.

Consider $(P_{Ij})_t = ((P_{Ij})_t^{12} \cdots (P_{Ij})_t^{lm} \cdots (P_{Ij})_t^{(k-1)k})$ where $(P_{Ij})_t^{lm}$ is of order $s \times 1$ and represents $(P_{Ij})_t$ for the two-factor interaction between lth and mth attribute. Therefore, the necessary and sufficient condition for $1'(P_{I1})_t = 1'(P_{I2})_t$ is that $1'(P_{I1})_t^{lm} = 1'(P_{I2})_t^{lm}$ for every l and m.

For the s choice pairs where either both attributes are having foldover in the second option or both are not having foldover in the second option, the corresponding rows in $(P_{I2})_t^{lm}$ is same as the corresponding rows in $(P_{I1})_t^{lm}$. However, for the pairs where one attribute has foldover in the second option and another does not have foldover in the second option, the corresponding rows in $(P_{I2})_t^{lm}$ is negative of the corresponding rows in $(P_{I1})_t^{lm}$.

Therefore to show $1'(P_{I1})_t^{lm} = 1'(P_{I2})_t^{lm}$ is equivalent to showing $1'(P_{I1})_t^{lm} = 0$ for the pairs where one attribute has foldover in the second option and another does not have foldover in the second option . Now, $1'(P_{I1})_t^{lm} = 0$ when the frequency of the pairs from the set $\{(01, 00), (01, 11), (10, 00), (10, 11)\}$ is same as the frequency of the pairs from set $\{(00, 01), (00, 10), (11, 01), (11, 10)\}$ under the lth and mth attribute.

The sufficiency condition follows trivially. ■

Note that a complete factorial design involving 2^k combinations can be equivalently written as the collection of 2^α combinations where each of the 2^α combinations have another replications of $2^{k-\alpha}$ combinations within each of them. We now give a method of construction to obtain an optimal paired choice block design for $\alpha \geq 2$ and $k-\alpha \geq 2$. The construction steps to arrive at a optimal paired choice block design d_2^I from d_1^I are:

(i) Make two sets of attributes positions F_F and F_{NF} respectively of size α and $k-\alpha$ representing attributes positions corresponding to attributes having foldover in the second option and attributes positions corresponding to attributes not having foldover in the second option

(ii) Write complete factorial for 2^α options. Note that this set can be divided into two-halves such that each element in first half has its foldover in the second half and vice versa. We consider any one set with $2^{\alpha-1}$ combinations.

(iii) Repeat step (ii) for $2^{k-\alpha}$ combinations.

(iv) Take one element from the first half of (ii) say a, put it in the first option in respective indices as in F_F. For the indices in F_{NF}, take any two different elements from (iii) say b and c. Find the respective foldovers from other halves say a', b' and c'. Without loss of generality say first α positions have foldover in the second option and rest $k-\alpha$ positions do not have foldover in the second option. Make the first block having 4 choice pairs as $\{ab, a'b\}$, $\{ab', a'b'\}$, $\{a'c, ac\}$, $\{a'c', a'c'\}$ and second block having choice pairs as $\{ac, a'c\}$, $\{ac', a'c'\}$, $\{a'b, ab\}$, $\{a'b', ab'\}$. Also, note that, the two constructed blocks satisfies the conditions for optimality in Theorem 5.1.

(v) Repeat step (iv) for each of $2^{\alpha-1}$ combinations of step (ii) using the same 2 elements from step (iii). Then, repeat the entire process for different 2 options from step (iii). This gives rise to a total set of $2^{\alpha-1}2^{k-\alpha-2} = 2^{k-3}$ set of two blocks of size 4 each.

Note that step (v) produces the complete factorial having 2^k combinations. Also, considering the fact that the two blocks created in step (iv) can be obtained from another by interchanging the options in the blocks, using any one block out of the two blocks is equally good. Thus, we have a total of 2^{k-1} choice pairs divided into 2^{k-3} blocks each of size 4.

(vi) Repeating steps (i)-(v) for α position changes, an optimal paired choice block design d_2^I is obtained with parameters $k, v=2, s=4, b$ where $b = 2^{k-3}\binom{k}{q}$ for k odd and $b = 2^{k-3}\binom{k+1}{q+1}$ for k even, and we have the following result.

Theorem 5.2. *For $k > 4$, there exists an optimal paired choice design d_1^I with parameters $k, v=2, s, b=1$ where $s = 2^{k-1}\binom{k}{q}$ for k odd and $s = 2^{k-1}\binom{k+1}{q+1}$ for k even. Furthermore, an optimal paired choice block design d_2^I with parameters $k, v=2, s=4, b$ is optimal in $\mathcal{D}_{k,b,s}$, where $b = 2^{k-3}\binom{k}{q}$ for k odd and $b = 2^{k-3}\binom{k+1}{q+1}$ for k even.*

We give an example to illustrate the construction.

Example 5.1. *Let $k=4, v=2, b=t_{10}, s=8$. Since $k=4$, our construction method does not allow to achieve d_2^I from d_1^I. However, for |alpha $=2$, the proposed construction method still holds. Thus we construct blocks of size 4 for $\alpha=2$ and for $\alpha=3$, we give a design with blocks of size 8.*

Starting with a 2^4 complete factorial design, a d_1^I exists with parameters $k=4, v=2, b=1, s=2^3(\binom{4}{2}+\binom{4}{3})=80$. For $\alpha=2$, foldovers in the second option are in positions

12,13,14,23,24 and 34 and for $\alpha = 3$*, foldovers in the second option are in positions 123,124,134 and 234. Consider* $\alpha = 2$ *and foldover in the second option in position say, 12. Let construction steps (ii) and (iii) have 00,01 elements in their first half and 11,10 in the second half. Then from construction step (iv), using 00 of step (ii) we form two blocks as* $\{(0000, 1100), (0011, 1111), (1101, 0001), (1110, 0010)\}$ *and* $\{(0001, 1101), (0010, 1110), (1100, 0000), (1110, 0011)\}$*. Since these two blocks are replica of each other, we consider only the first block. Proceeding in this way, we get the design as below for* $\alpha = 2$*.*

B_1		B_2		B_3		B_4	
(0000)	(1100)	(0100)	(1000)	(0000)	(1010)	(0010)	(1000)
(0011)	(1111)	(0111)	(1011)	(0101)	(1111)	(0111)	(1101)
(1101)	(0001)	(1001)	(0101)	(1011)	(0001)	(1001)	(0011)
(1110)	(0010)	(1010)	(0110)	(1110)	(0100)	(1100)	(0110)
B_5		B_6		B_7		B_8	
(0000)	(1001)	(0001)	(1000)	(0000)	(0110)	(0010)	(0100)
(0110)	(1111)	(0111)	(1110)	(1001)	(1111)	(1011)	(1101)
(1011)	(0010)	(1010)	(0011)	(0111)	(0001)	(0101)	(0011)
(1101)	(0100)	(1100)	(0101)	(1110)	(1000)	(1100)	(1010)
B_9		B_{10}		B_{11}		B_{12}	
(0000)	(0101)	(0001)	(0100)	(0000)	(0011)	(0001)	(0010)
(0110)	(0011)	(0111)	(0010)	(1100)	(1111)	(1101)	(0010)
(1011)	(1110)	(1010)	(1111)	(0111)	(0100)	(0110)	(0101)
(1101)	(1000)	(1100)	(1001)	(1011)	(1000)	(1010)	(1001)

For $\alpha = 3$*, since* $k - \alpha = 1$*, the above procedure does not work and we provide a design in 4 blocks each of size 8.*

B_{13}		B_{14}		B_{15}		B_{16}	
(0000)	(1110)	(0000)	(1011)	(0000)	(1101)	(0000)	(0111)
(0110)	(1000)	(0011)	(1000)	(0101)	(1000)	(0011)	(0100)
(1010)	(0100)	(1010)	(0001)	(1100)	(0001)	(0101)	(0010)
(1100)	(0010)	(1001)	(0010)	(1001)	(0100)	(0110)	(0001)
(0001)	(1111)	(0100)	(1111)	(0010)	(1111)	(1000)	(1111)
(0111)	(1001)	(0111)	(1100)	(0111)	(1010)	(1011)	(1100)
(1011)	(0101)	(1110)	(0101)	(1110)	(0011)	(1101)	(1010)
(1101)	(0011)	(1101)	(0110)	(1011)	(0110)	(1110)	(1001)

Combining two blocks of size 4 each for $\alpha = 2$*, we get 6 blocks of size 8, which in combination with the 4 blocks of size 8 corresponding to* $\alpha = 3$ *gives the desired optimal design.*

6 Concluding remarks

For given k and v_i, we have considered designs for choice experiments where each respondent is shown s choice pairs. However, the optimal designs obtained in this paper for various block sizes can be clubbed to get an optimal paired choice block design with unequal block sizes. For example, under the main effects model, let $X = (X_1' \cdots X_b')'$. Also, let the block size of the tth block be s_t. Then, the information matrix $\tilde{C}_M = C_M - \sum_{t=1}^{b}(1/s_t)X_t' 1_t 1_t' X_t$, and Theorem 3.1 still holds. Thus, the choice design obtained by combining w optimal paired choice block designs with parameters $k, v_1, \ldots, v_k, b_i, s_i; i = 1, \ldots, w$ is optimal. Similar thing also holds for optimal paired choice designs with unequal block sizes under the main plus two-factor interaction effects model.

References

[1] Michiel CJ Bliemer and John Matthew Rose. Efficiency and sample size requirements for stated choice experiments. In *Transportation Research Board 88th Annual Meeting*, number 09-2421, 2009.

[2] Stephen Bush. Optimal designs for stated choice experiments generated from fractional factorial designs. *J. Stat. Theory Pract.*, 8(2):367–381, 2014.

[3] Stephen Bush, Deborah J Street, and Leonie Burgess. Optimal designs for stated choice experiments that incorporate position effects. *Comm. Statist. Theory Methods*, 41(10):1771–1795, 2012.

[4] Fatih Demirkale, Diane Donovan, and Deborah J Street. Constructing D-optimal symmetric stated preference discrete choice experiments. *J. Statist. Plann. Inference*, 143(8):1380–1391, 2013.

[5] Peter Goos and Heiko Großmann. Optimal design of factorial paired comparison experiments in the presence of within-pair order effects. *Food quality and preference*, 22(2):198–204, 2011.

[6] Ulrike Graßhoff, Heiko Großmann, Heinz Holling, and Rainer Schwabe. Optimal paired comparison designs for first-order interactions. *Statistics*, 37(5):373–386, 2003.

[7] Ulrike Graßhoff, Heiko Großmann, Heinz Holling, and Rainer Schwabe. Optimal designs for main effects in linear paired comparison models. *J. Statist. Plann. Inference*, 126(1):361–376, 2004.

[8] Heiko Großmann, Heinz Holling, and Rainer Schwabe. Advances in optimum experimental design for conjoint analysis and discrete choice models. In P H Franses and A L Montgomery, editors, *Econometric Models in Marketing. Advances in Econometrics*, volume 16, pages 93–118. JAI Press INC., Amsterdam, 2002.

[9] Heiko Großmann and Rainer Schwabe. Design for discrete choice experiments. In A Dean, M Morris, J Stufken, and D. Bingham, editors, *Handbook of Design and Analysis of Experiments*, pages 791–835. Chapman and Hall/CRC, Boca Raton, 2015.

[10] A. S. Hedayat, N. J. A. Sloane, and John Stufken. *Orthogonal arrays*. Springer Series in Statistics. Springer-Verlag, New York, 1999. Theory and applications, With a foreword by C. R. Rao.

[11] Stephane Hess and John M Rose. Some lessons in stated choice survey design. In *European Transport Conference, 2009*, 2009.

[12] Joel Huber and Klaus Zwerina. The importance of utility balance in efficient choice designs. *J. Mark. Res.*, 33:307–317, 1996.

[13] Rakhi Singh, Feng-Shun Chai, and Ashish Das. Optimal two-level choice designs for any number of choice sets. *Biometrika*, to appear, 2015.

[14] Deborah J. Street, David S. Bunch, and Beverley J. Moore. Optimal designs for 2^k paired comparison experiments. *Comm. Statist. Theory Methods*, 30(10):2149–2171, 2001.

[15] Deborah J. Street and Leonie Burgess. *The construction of optimal stated choice experiments: Theory and methods*, volume 647. John Wiley & Sons, Hoboken, New Jersey, 2007.