Optimal two-level conjoint designs with constant attributes in the profile sets

Roselinde Kessels [a,1], Peter Goos [a,b,*], Martina Vandebroek [c,d,2]

[a] *Universiteit Antwerpen, Faculty of Applied Economics and StatUa Center for Statistics, Prinsstraat 13, 2000 Antwerpen, Belgium*
[b] *Erasmus Universiteit Rotterdam, Erasmus School of Economics, Postbus 1738, 3000 DR Rotterdam, The Netherlands*
[c] *Katholieke Universiteit Leuven, Faculty of Business and Economics, Naamsestraat 69, 3000 Leuven, Belgium*
[d] *Leuven Statistics Research Centre, Celestijnenlaan 200B, 3001 Leuven-Heverlee, Belgium*

ARTICLE INFO

ABSTRACT

We propose a simple strategy to construct D-, A-, G- and V-optimal two-level designs for rating-based conjoint studies with large numbers of attributes. In order to simplify the rating task, the designs hold one or more attributes at a constant level in each profile set. Our approach combines orthogonal designs and binary incomplete block designs with equal replication. The designs are variance-balanced meaning that they yield an equal amount of information on each of the part-worths.

© 2010 Published by Elsevier B.V.

1. Introduction

A conjoint experiment tries to elucidate consumer preferences for the attributes of a good, that is, a product or a service. This can be done by asking respondents to rate a set of alternative goods on a scale or to give the price they are willing to pay for a certain good. Each alternative or profile involves a combination of levels of a set of predefined attributes. The objective of a conjoint experiment is to elicit as much information as possible on the utilities people derive from the attribute levels. These utilities are also called part-worths and correspond to the parameters of the statistical model that is used to analyze the data. Ultimately, companies can use the information gained to make precise predictions of consumers' preferences, and thereby develop new goods that lead to a substantial rise in market share.

Reaching this goal requires precise estimates of the parameters in the statistical model. Constructing efficient conjoint designs is key, not only for precise predictions but also for reducing the cost of the experiment. The conjoint designs we set up in this paper are intended for screening the vital few important attributes from a group of many potential ones. The different attributes involved in the conjoint designs have two levels each. In most conjoint studies, see e.g. Danaher

(1997), Pullman et al. (2002) and references therein, the conjoint design consists of blocks or sets of an equal number of profiles. Each of these sets is evaluated by a different respondent.

The incorporation of large numbers of attributes in the design of conjoint experiments requires some special attention. As Green (1974) and Schwabe et al. (2003) have argued, respondents may get overloaded when they have to assimilate profiles that embrace more than four different attribute levels. Therefore, to overcome respondent fatigue, it is advised to vary only the levels of a small number of attributes in each of the profile sets assigned to the respondents. The profiles are still combinations of levels for all attributes, but the levels of one or more attributes do not vary in a profile set. Each respondent thus rates a set of profiles in which the levels of one or more attributes are held constant. These constant attributes need not be the same in each profile set.

Because each of the profile sets has a number of attributes at a constant level, the conjoint designs addressed in this paper show some similarities to split-plot designs. Split-plot designs are heavily used in industry when the levels of some of the experimental factors are difficult or costly to change or control. These factors are called whole-plot factors and are kept at a constant level for several observations in the design. Split-plot designs are similar to the type of conjoint design that we study here because they also involve sets of observations within which some of the factors' levels, the whole-plot factors' levels, are held constant. Split-plot designs differ from our conjoint design in that the factors whose levels are held constant are naturally the same in each set of the split-plot design, whereas the constant attributes may change from set to set in the conjoint design. We refer to the work of Goos and Vandebroek (2001, 2003, 2004) and Goos (2002, 2006) for more details on split-plot designs.

Our approach to construct multi-attribute conjoint designs is based on the same linear random effects model that is used for generating split-plot designs. In our conjoint setting, the random effects are attributable to a respondent who rates the profile set. The reason is that respondents are heterogeneous, meaning that profile ratings from the same respondent are more similar than profile ratings from different respondents. Also Brazier et al. (2002) and Kessels et al. (2008) adopted this model for conjoint applications. Because the conjoint designs in this paper are set up for identifying the most important attributes from a large number of attributes, only main-effects models are considered.

Our design construction approach is conceptually easy to understand and generates designs with sets of 2, 4 and 8 profiles. In principle, our approach can produce designs whenever the size of the profile sets is a power of two. Related to our conjoint designs with sets of two profiles are paired comparison designs (Grossmann et al., 2002, 2005; Grasshoff et al., 2004; Street et al., 2001; Street and Burgess, 2004). These designs also comprise sets of two profiles, but in contrast with conjoint designs, each respondent evaluates all the sets. This is done by specifying the preferred profile in each set, and possibly also the preference strength.

Street et al. (2001) demonstrated that holding the levels of one or more attributes constant in two-level paired comparison designs leads to information losses when main-effects models are considered. Only when fold-over pairs, in which the signs of the attribute levels of the first profile are reversed in the second profile, are used, the information on the part-worths is maximized. This result also applies to conjoint designs in which the respondents rate two profiles each. In general, keeping the rating task of a number of profiles manageable for the respondents through the use of constant attributes comes at a loss of information on the part-worths. Since all part-worths are assumed to be on the same footing, we look for conjoint designs that provide an equal amount of information on each part-worth. We call these designs variance-balanced.

To allow for variance balance in the conjoint designs, the constant attributes have to differ between the profile sets such that each attribute is constant in an equal number of sets. For problems with more than one constant attribute, we obtain an appropriate pattern of constant attributes using a binary incomplete block design with equal replication of all treatments (e.g. a balanced or partially balanced incomplete block design (BIBD or PBIBD)). The equal replication of the treatments ensures that every attribute is held constant equally often and paves the way for the production of variance-balanced conjoint designs. To ensure the optimality of the designs, meaning that they yield maximum information on each part-worth, we choose the design profiles so that they form an orthogonal array. These designs have columns that are pairwise orthogonal and therefore yield parameter estimates that are statistically independent of each other.

The outline of the remainder of the paper is as follows. Section 2 reviews the random respondent effects model used in conjoint design. In Section 3 we discuss the conditions needed to obtain universally optimal variance-balanced designs. We explain our design construction approach in Section 4 and describe the information content of the resulting designs in Section 5. Section 6 concludes the paper and highlights some further research possibilities.

2. The random respondent effects model

The model used to set up and analyze two-level multi-attribute conjoint experiments is the random respondent effects model. In this model, it is assumed that respondents are heterogeneous and randomly selected from a prespecified population. Each respondent i, $i=1,\ldots,b$, rates a different block or set of profiles to estimate all parameters. For convenience, the profile sets assigned to the b respondents have the same size m. As a result, the total number of experimental profiles amounts to $n=bm$. In this paper, we propose designs for two-level conjoint studies where m equals either 2, 4 or 8. The rating U_{ij} for profile j, $j=1,\ldots,m$, by a respondent i is modelled as

$$U_{ij} = \mathbf{x}'_{ij}\boldsymbol{\beta} + \gamma_i + \varepsilon_{ij}. \tag{1}$$

In the model, $\mathbf{x}_{ij}$ is a $(k+1) \times 1$ vector that has a one as first element (corresponding to the intercept) and the attribute levels describing the jth profile rated by respondent i as the remaining k elements. The number k is the number of attributes in the study. Each of these attributes can be quantitative or categorical. We use a -1 and a $+1$ to code the two levels of each attribute. The vector $\boldsymbol{\beta} = [\beta_0, \beta_1, \ldots, \beta_k]'$ contains the $(k+1)$ unknown parameters with β_0 the intercept and $\beta_1, \ldots, \beta_k$ the part-worths or weights attached to the attribute levels. The term γ_i represents the random effect of respondent i, whereas ε_{ij} is a random error term.

In matrix notation, model (1) can be written as

$$\boldsymbol{U} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \tag{2}$$

where $\boldsymbol{U}$ is a vector of n profile ratings, the vector $\boldsymbol{\gamma} = [\gamma_1, \ldots, \gamma_b]'$ contains the b random respondent effects and $\boldsymbol{\varepsilon}$ is a random error vector. The matrices $\mathbf{X}$ and $\mathbf{Z}$ have dimensions $n \times (k+1)$ and $n \times b$, respectively. The model matrix $\mathbf{X}$ is given by $[\mathbf{X}_1', \ldots, \mathbf{X}_b']'$, where $\mathbf{X}_i = [\mathbf{x}_{i1}, \ldots, \mathbf{x}_{im}]'$ contains the attribute levels of the m profiles rated by respondent i, whereas $\mathbf{Z} = \mathbf{I}_b \otimes \mathbf{1}_m$, with $\otimes$ the Kronecker product and $\mathbf{1}_m$ an $m \times 1$ vector of ones. It is assumed that

$$\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}_n \quad \text{and} \quad \mathrm{Cov}(\boldsymbol{\varepsilon}) = \sigma_\varepsilon^2 \mathbf{I}_n, \tag{3}$$

$$\mathrm{E}(\boldsymbol{\gamma}) = \mathbf{0}_b \quad \text{and} \quad \mathrm{Cov}(\boldsymbol{\gamma}) = \sigma_\gamma^2 \mathbf{I}_b, \tag{4}$$

$$\mathrm{Cov}(\boldsymbol{\gamma}, \boldsymbol{\varepsilon}) = \mathbf{0}_{b \times n}, \tag{5}$$

where σ_ε^2 is the variance within respondents and σ_γ^2 is the variance between respondents. Under these assumptions, the variance–covariance matrix $\mathbf{V}$ of the profile ratings $\boldsymbol{U}$ is $\mathbf{I}_b \otimes \mathbf{V}_m$, with

$$\mathbf{V}_m = \sigma_\varepsilon^2 \mathbf{I}_m + \sigma_\gamma^2 \mathbf{1}_m \mathbf{1}_m' = \sigma_\varepsilon^2 \left(\mathbf{I}_m + \frac{\rho}{1-\rho} \mathbf{1}_m \mathbf{1}_m' \right). \tag{6}$$

In this expression, $\rho = \sigma_\gamma^2 / (\sigma_\varepsilon^2 + \sigma_\gamma^2)$ measures the proportion of the total variance that is accounted for by the differences between respondents. It represents the degree of correlation between the ratings from a single respondent.

The vector of the unknown fixed model parameters $\boldsymbol{\beta}$ can be estimated using the generalized least squares estimator

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\boldsymbol{U}, \tag{7}$$

with $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$ the variance–covariance matrix of $\hat{\boldsymbol{\beta}}$, the inverse of which is the information matrix on $\boldsymbol{\beta}$.

$\mathcal{D}$-optimal designs for the random respondent effects model minimize the determinant of the variance–covariance matrix $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$, or, equivalently, maximize the determinant of the information matrix $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$. $\mathcal{A}$-optimal designs minimize the average variance of the parameter estimates and thus the trace of $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$. Finally, $\mathcal{G}$- and $\mathcal{V}$-optimal designs minimize the maximum and average prediction variance $\mathbf{x}_{ij}'(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{x}_{ij}$, respectively.

3. Optimal variance-balanced designs

In order not to overload the respondents with a heavy rating task, we hold one or more attributes at a fixed level in each of the profile sets. In other words, we introduce perfect level overlap for one or more attributes in each profile set. These attributes are the *constant attributes* and may differ from set to set. We denote the number of constant attributes by k_c. The remaining $k_v = k - k_c$ attributes, the levels of which may vary within a profile set, are the *non-constant attributes*. The number k_v is similar to the comparison depth used by Grasshoff et al. (2003) in the context of paired comparison experiments. The concepts of constant and non-constant attributes are illustrated by means of the conjoint design in Table 1. This design consists of $b = 12$ sets with $m = 2$ profiles each. In every profile set, the levels of two of the six attributes are constant while the levels of the other four attributes change from profile to profile. Hence $k = 6$, $k_c = 2$ and $k_v = 4$. The levels of the constant attributes are in bold.

The design strategy we propose allows for the generation of conjoint designs that provide a maximum and equal amount of information on each part-worth under the constraint that the levels of k_c attributes are constant within each profile set. As such, we assume that all part-worths are equally important. Because the designs we propose allow every part-worth to be estimated with the same precision, they are called variance-balanced. The property of variance-balance of the designs as well as their $\mathcal{D}$-, $\mathcal{A}$-, $\mathcal{G}$- and $\mathcal{V}$-optimality are proven in the following theorem.

Theorem. *The information matrix of a $\mathcal{D}$-, $\mathcal{A}$-, $\mathcal{G}$- and $\mathcal{V}$-optimal variance-balanced two-level rating-based conjoint design with b profile sets of size m and with k_c of the k attributes held constant in each profile set is given by*

$$\mathbf{X}'\mathbf{V}^{-1}\mathbf{X} = \frac{bm}{\sigma_\varepsilon^2} \operatorname{diag}\left[c_m; \frac{k_c}{k} c_m + \frac{k_v}{k}; \frac{k_c}{k} c_m + \frac{k_v}{k}; \ldots; \frac{k_c}{k} c_m + \frac{k_v}{k} \right], \tag{8}$$

where $k_v = k - k_c$ and $c_m = (1-\rho)/(1+(m-1)\rho)$. This information matrix can be obtained by choosing the levels of the constant and non-constant attributes such that:

1. *the levels of the constant attributes in the profile sets with the same constant attributes are given by an orthogonal array,*
2. *the levels of the non-constant attributes in the profile sets with the same constant attributes are given by an orthogonal array,*

Table 1
Optimal variance-balanced conjoint design with $b=12$ sets of $m=2$ profiles, $k_c=2$ constant attributes and $k_v=4$ non-constant attributes.

Set	Attributes					
	1	2	3	4	5	6
1	**−1**	**−1**	−1	−1	−1	−1
1	**−1**	**−1**	+1	+1	+1	+1
2	**−1**	**+1**	−1	+1	−1	+1
2	**−1**	**+1**	+1	−1	+1	−1
3	**+1**	**−1**	−1	−1	+1	+1
3	**+1**	**−1**	+1	+1	−1	−1
4	**+1**	**+1**	−1	+1	+1	−1
4	**+1**	**+1**	+1	−1	−1	+1
5	−1	−1	**−1**	**−1**	−1	+1
5	+1	+1	**−1**	**−1**	+1	−1
6	−1	+1	**−1**	**+1**	−1	−1
6	+1	−1	**−1**	**+1**	+1	+1
7	−1	−1	**+1**	**−1**	+1	−1
7	+1	+1	**+1**	**−1**	−1	+1
8	−1	+1	**+1**	**+1**	+1	+1
8	+1	−1	**+1**	**+1**	−1	−1
9	−1	−1	−1	−1	**−1**	**−1**
9	+1	+1	+1	+1	**−1**	**−1**
10	−1	+1	−1	+1	**−1**	**+1**
10	+1	−1	+1	−1	**−1**	**+1**
11	−1	−1	+1	+1	**+1**	**−1**
11	+1	+1	−1	−1	**+1**	**−1**
12	−1	+1	+1	−1	**+1**	**+1**
12	+1	−1	−1	+1	**+1**	**+1**

3. *the levels of the constant attributes sum to zero over the profile sets in which the same attributes are kept constant,*
4. *the levels of the non-constant attributes sum to zero in each of the b profile sets, and*
5. *each attribute is constant in an equal number of profile sets.*

Proof. To prove the theorem, we show that any design satisfying the requirements 1–4 has a diagonal information matrix, with diagonal elements that are maximal. It follows from the lemma in the Appendix to this paper that the designs are therefore $\mathcal{D}$-, $\mathcal{A}$-, $\mathcal{G}$- and $\mathcal{V}$-optimal. We now show that a design that satisfies requirements 1–4 has maximum diagonal elements. Next, we explain that if, in addition, requirement 5 is satisfied, the design is also variance-balanced.

Goos and Vandebroek (2001) showed that, under assumptions (3), (4) and (5), $\mathbf{V}$ is block diagonal, and

$$\mathbf{X}'\mathbf{V}^{-1}\mathbf{X} = \sum_{i=1}^{b} \mathbf{X}_i'\mathbf{V}_m^{-1}\mathbf{X}_i, \tag{9}$$

where

$$\mathbf{V}_m^{-1} = \sigma_\varepsilon^{-2}\left(\mathbf{I}_m - \frac{\rho}{1+\rho(m-1)}\mathbf{1}_m\mathbf{1}_m'\right). \tag{10}$$

According to (9), the information matrix of a conjoint design is the sum of b information matrices, one for each of the profile sets. Now, consider a given profile set i with design matrix $\mathbf{X}_i$, and assume for notational convenience that the attributes have been rearranged so that the first k_c attributes are constant. In that case we can partition $\mathbf{X}_i$ as $[\mathbf{1}_m\ \mathbf{1}_m\mathbf{w}_i'\ \mathbf{S}_i]$ where $\mathbf{w}_i$ is the $k_c \times 1$ vector containing the levels of the k_c constant attributes and $\mathbf{S}_i$ is an $m \times k_v$ matrix containing the levels of the k_v non-constant attributes in the ith profile set. The information matrix corresponding to the ith profile set then is of the form

$$\mathbf{X}_i'\mathbf{V}_m^{-1}\mathbf{X}_i = \begin{pmatrix} \mathbf{1}_m'\mathbf{V}_m^{-1}\mathbf{1}_m & \mathbf{1}_m'\mathbf{V}_m^{-1}\mathbf{1}_m\mathbf{w}_i' & \mathbf{1}_m'\mathbf{V}_m^{-1}\mathbf{S}_i \\ \mathbf{w}_i\mathbf{1}_m'\mathbf{V}_m^{-1}\mathbf{1}_m & \mathbf{w}_i\mathbf{1}_m'\mathbf{V}_m^{-1}\mathbf{1}_m\mathbf{w}_i' & \mathbf{w}_i\mathbf{1}_m'\mathbf{V}_m^{-1}\mathbf{S}_i \\ \mathbf{S}_i'\mathbf{V}_m^{-1}\mathbf{1}_m & \mathbf{S}_i'\mathbf{V}_m^{-1}\mathbf{1}_m\mathbf{w}_i' & \mathbf{S}_i'\mathbf{V}_m^{-1}\mathbf{S}_i \end{pmatrix}. \tag{11}$$

Using (10), this can be simplified to

$$
\mathbf{X}_i' \mathbf{V}_m^{-1} \mathbf{X}_i = \sigma_\varepsilon^{-2}
\begin{pmatrix}
mc_m & mc_m \mathbf{w}_i' & c_m \mathbf{1}_m' \mathbf{S}_i \\
mc_m \mathbf{w}_i & mc_m \mathbf{w}_i \mathbf{w}_i' & mc_m \mathbf{w}_i \mathbf{1}_m' \mathbf{S}_i \\
c_m \mathbf{S}_i' \mathbf{1}_m & c_m \mathbf{S}_i' \mathbf{1}_m \mathbf{w}_i' & \mathbf{S}_i' \mathbf{S}_i - c_m \dfrac{\rho}{1-\rho} (\mathbf{S}_i' \mathbf{1}_m)(\mathbf{S}_i' \mathbf{1}_m)'
\end{pmatrix}.
\tag{12}
$$

The first diagonal element, mc_m, is a constant which does not depend on the choice of the design. The next k_c diagonal elements, which correspond to the constant attributes, are also equal to mc_m. This is because for any two-level design whose first k_c attributes are constant in profile set i, $\mathbf{w}_i$ is a row of -1's and $+1$'s so that each diagonal element of $\mathbf{w}_i \mathbf{w}_i'$ is one. For any two-level design $\mathbf{S}_i$, the diagonal elements of $\mathbf{S}_i' \mathbf{S}_i$ are equal to m, and if requirement 4 is satisfied, $\mathbf{S}_i' \mathbf{1}_m = \mathbf{0}_{k_v}$, so that the k_v diagonal elements corresponding to the non-constant attributes are equal to m. Requirement 4 therefore ensures that the diagonal elements of the information matrix for profile set i are maximal.

Now consider all g profiles sets for which the first k_c attributes are constant, and which therefore have information matrices similar to that in (12). Each of these matrices has maximum diagonal elements, and, hence, their sum also has maximum diagonal elements. Moreover, if requirements 1–3 are satisfied, then $\sum_{i=1}^{g} \mathbf{w}_i \mathbf{w}_i' = g \mathbf{I}_{k_c}$, $\sum_{i=1}^{g} \mathbf{w}_i = \mathbf{0}_{k_c}$ and $\sum_{i=1}^{g} \mathbf{S}_i' \mathbf{S}_i = mg \mathbf{I}_{k_v}$. As a result,

$$
\sum_{i=1}^{g} \mathbf{X}_i' \mathbf{V}_m^{-1} \mathbf{X}_i = \sigma_\varepsilon^{-2}
\begin{pmatrix}
mgc_m & \mathbf{0}_{k_c}' & \mathbf{0}_{k_v}' \\
\mathbf{0}_{k_c} & mgc_m \mathbf{I}_{k_c} & \mathbf{0}_{k_c \times k_v} \\
\mathbf{0}_{k_v} & \mathbf{0}_{k_v \times k_c} & mg \mathbf{I}_{k_v}
\end{pmatrix}.
\tag{13}
$$

Summing the information matrices of all g profiles sets in which the first k_c attributes are constant thus produces a diagonal information matrix with maximum diagonal elements. In a similar fashion, we can derive the information matrix corresponding to any other set of g profile sets in which another set of attributes is held constant. If requirement 5 is satisfied for the entire design, then there are b/g such sets of g profile sets and each attribute is constant in bk_c/k profile sets. As a result, the total information matrix of the entire design will equal (8). Because the diagonal elements corresponding to the $k = k_c + k_v$ attributes are all equal and the matrix is diagonal, every part-worth estimate has the same variance, so that any design satisfying requirements 1–5 is variance-balanced. As shown in the Appendix to this paper, the fact that the diagonal elements of the total information matrix are maximal ensures that the designs are $\mathcal{D}$-, $\mathcal{A}$-, $\mathcal{G}$- and $\mathcal{V}$-optimal. $\square$

4. Design construction approach

In this section, we illustrate our strategy to set up optimal two-level variance-balanced conjoint designs in which a given number of attributes is constant in each profile set. While our construction approach is very general and can be used for large numbers of attributes, we focus on designs with small numbers of attributes for the clarity of the exposition. For a catalog of designs with larger numbers of attributes, we refer the reader to Kessels et al. (2006). First, we consider situations where $k_c=1$, and next, we discuss cases where $k_c > 1$. To develop the optimal variance-balanced conjoint designs, we need to determine:

1. the constant attributes in each profile set,
2. the levels of these constant attributes in each profile set, and
3. the levels of the non-constant attributes in each profile set.

4.1. Optimal variance-balanced conjoint designs for $k_c=1$

This section shows how to construct optimal variance-balanced conjoint designs with $k_c=1$ constant attribute, $k_v=2$, 3 or 4 non-constant attributes and $m=2$, 4 or 8 profiles per set. If $k_c=1$, determining the constant attributes and their levels is straightforward. To ensure variance balance, each attribute should be constant in an equal number of profile sets. That number of profile sets has to be even so that the number of sets in which the attributes are held constant at level -1 matches the number of sets in which they are held at level $+1$. To determine the levels of the non-constant attributes, we use small orthogonal arrays as subdesigns. For $k_v=2$, 3 or 4 non-constant attributes and $m=2$, 4 or 8 profiles per set, we construct the smallest possible orthogonal subdesigns with an even number of profile sets and match these with the constant attributes to produce optimal variance-balanced conjoint designs. We discuss the use of orthogonal subdesigns for each case.

The orthogonal subdesigns shown in Tables 2a and b determine the levels of $k_v=2$ non-constant attributes for profile sets of sizes $m=2$ and 4, respectively. The complete subdesign for $m=2$ as well as each of the profile sets in the subdesign for $m=4$ represents the full 2^2 factorial design. Both subdesigns consist of two profile sets and satisfy requirements 2 and 4 of the theorem.

Table 2
Orthogonal subdesigns with the levels of $k_v=2$ non-constant attributes.

Set	Non-constant attributes	
	1	2
(a) $m=2$		
1	−1	−1
1	+1	+1
2	−1	+1
2	+1	−1
(b) $m=4$		
1	−1	−1
1	+1	+1
1	−1	+1
1	+1	−1
2	−1	−1
2	+1	+1
2	−1	+1
2	+1	−1

To create optimal variance-balanced conjoint designs for $k_c=1$ and profile sets of sizes $m=2$ or 4, the appropriate subdesign from Tables 2a or b should be incorporated at least three times. This is the required minimum to have each of the $k=3$ attributes act as a constant attribute. This yields $b=2k=6$ profile sets which can be replicated if necessary. Note that it is not realistic to construct an optimal variance-balanced conjoint design for $k_v=2$ non-constant attributes and sets of $m=8$ profiles. This is because the profile sets in such design would contain every profile twice which makes no sense in a conjoint study.

Tables 3a, b and c contain the orthogonal subdesigns with the levels for $k_v=3$ non-constant attributes. The subdesign of Table 3a embraces four profile sets of size $m=2$ and the subdesigns of Tables 3b and c embrace two profile sets of sizes $m=4$ and 8. The full 2^3 factorial design was exploited to build the subdesigns. It is included once in the subdesigns in Tables 3a and b and twice in the subdesign in Table 3c.

Each of the subdesigns in Tables 3a, b and c should be incorporated at least four times to build optimal variance-balanced conjoint designs for $k_c=1$. This allows each of the $k=4$ attributes to be constant in at least two profile sets. The smallest conjoint design for $m=2$ then involves 16 profile sets in total, while the smallest designs for $m=4$ and 8 require eight profile sets.

The last orthogonal subdesigns we discuss appear in Tables 4a, b and c, and involve $k_v=4$ non-constant attributes. Table 4a shows two possible subdesigns with four profile sets of size $m=2$ and Table 4b shows two possible subdesigns with two profile sets of size $m=4$. The subdesign in Table 4c has two profile sets of size $m=8$. The combinations of the two possible subdesigns in Tables 4a and b as well as the subdesign in Table 4c form the full 2^4 factorial design. The subdesigns in Tables 4a and b each constitute an orthogonal fraction of this design.

The smallest optimal variance-balanced conjoint designs require five subdesigns to allow each of the $k=5$ attributes to act as a constant attribute. If $m=2$ or 4 it does not matter whether subdesign (i) or (ii) is selected from Tables 4a or b. Various replication structures of the subdesigns are possible because the optimality of the designs is not affected by the choice of subdesigns used. The smallest conjoint design for $m=2$ requires 20 profile sets, while the smallest designs for $m=4$ and 8 involve 10 profile sets.

In a similar fashion it is possible to create conjoint experiments when the number of non-constant attributes k_v is larger than four.

4.2. Optimal variance-balanced conjoint designs for $k_c > 1$

In this section, we extend our design strategy to allow for the creation of optimal variance-balanced conjoint designs with $k_c > 1$ constant attributes, $k_v=2$, 3 or 4 non-constant attributes and $m=2$, 4 or 8 profiles per set. Designs with $k_v > 4$ can be developed analogously.

4.2.1. Determining the constant attributes

If $k_c > 1$, it may be quite cumbersome to look for combinations of constant attributes that satisfy the requirements 1, 3 and 5 of the theorem. This is particularly true when k_c is large and not a divisor of k. Therefore, we use binary incomplete block designs to determine patterns of constant attributes that result in variance-balanced conjoint designs. Binary incomplete block designs describe how to arrange the levels of a single qualitative factor, called treatments, in groups or blocks of a certain size. Every treatment appears at most once in every block of a binary incomplete block design. What is crucial for the construction of our variance-balanced designs is that the total number of times a treatment appears in the

Table 3
Orthogonal subdesigns with the levels of $k_v = 3$ non-constant attributes.

Set	Non-constant attributes		
	1	2	3
(a) $m=2$			
1	−1	−1	−1
1	+1	+1	+1
2	−1	−1	+1
2	+1	+1	−1
3	−1	+1	−1
3	+1	−1	+1
4	−1	+1	+1
4	+1	−1	−1
(b) $m=4$			
1	−1	−1	−1
1	+1	+1	+1
1	−1	−1	+1
1	+1	+1	−1
2	−1	+1	−1
2	+1	−1	+1
2	−1	+1	+1
2	+1	−1	−1
(c) $m=8$			
1	−1	−1	−1
1	+1	+1	+1
1	−1	−1	+1
1	+1	+1	−1
1	−1	+1	−1
1	+1	−1	+1
1	−1	+1	+1
1	+1	−1	−1
2	−1	−1	−1
2	+1	+1	+1
2	−1	−1	+1
2	+1	+1	−1
2	−1	+1	−1
2	+1	−1	+1
2	−1	+1	+1
2	+1	−1	−1

binary incomplete block design is the same for every treatment. This is because the equal replication of the treatments ensures that every attribute in our conjoint designs acts as a constant attribute equally often. Well-known binary incomplete block designs that have equal replication of all treatments are balanced and partially balanced incomplete block designs (BIBDs and PBIBDs). We refer to Cochran and Cox (1957) and Cox (1958) for a general account of BIBDs and PBIBDs and to Abel and Greig (2006) and Mathon and Rosa (2006) for more discussions and tables of BIBDs. PBIBDs are fully elaborated in Shah and Sinha (1989) and Street and Street (1996), and catalogs of these designs can be found in Clatworthy (1973) and on a website by Sinha (see the list of references). There are, however, many binary incomplete block designs other than BIBDs and PBIBDs in which each treatment is repeated equally often and, hence, can be used as building blocks for our variance-balanced designs.

As an example, we refer to the conjoint design in Table 1 which we created using a binary incomplete block design with blocks {1,2}, {3,4} and {5,6}. This design has three blocks of size two, indicating that three types of profile sets are required. The first type has attributes 1 and 2 as constant attributes. The second type has attributes 3 and 4 as constant attributes, and the third type requires attributes 5 and 6 to be constant. In Table 1, there are four instances of every type of profile set. How to determine that number of instances is explained in the next section.

4.2.2. Determining the levels of the constant and non-constant attributes

We complete the description of our design methodology for $k_c > 1$ constant attributes by addressing the questions at what levels the constant attributes have to be set, how often they have to be held fixed and what levels the non-constant attributes have to take.

First, to determine the levels of $k_c > 1$ constant attributes, we use either a full 2^{k_c} factorial design or an orthogonal fraction of it. Within every profile set, we replicate one combination of attribute levels m times for the k_c constant

Table 4
Orthogonal subdesigns with the levels of $k_v=4$ non-constant attributes.

Set	Non-constant attributes			
	1	2	3	4
(a) $m=2$				
(i)				
1	-1	-1	-1	-1
1	$+1$	$+1$	$+1$	$+1$
2	-1	-1	$+1$	$+1$
2	$+1$	$+1$	-1	-1
3	-1	$+1$	-1	$+1$
3	$+1$	-1	$+1$	-1
4	-1	$+1$	$+1$	-1
4	$+1$	-1	-1	$+1$
(ii)				
1	-1	-1	-1	$+1$
1	$+1$	$+1$	$+1$	-1
2	-1	-1	$+1$	-1
2	$+1$	$+1$	-1	$+1$
3	-1	$+1$	-1	-1
3	$+1$	-1	$+1$	$+1$
4	-1	$+1$	$+1$	$+1$
4	$+1$	-1	-1	-1
(b) $m=4$				
(i)				
1	-1	-1	-1	-1
1	$+1$	$+1$	$+1$	$+1$
1	-1	-1	$+1$	$+1$
1	$+1$	$+1$	-1	-1
2	-1	$+1$	-1	$+1$
2	$+1$	-1	$+1$	-1
2	-1	$+1$	$+1$	-1
2	$+1$	-1	-1	$+1$
(ii)				
1	-1	-1	-1	$+1$
1	$+1$	$+1$	$+1$	-1
1	-1	-1	$+1$	-1
1	$+1$	$+1$	-1	$+1$
2	-1	$+1$	-1	-1
2	$+1$	-1	$+1$	$+1$
2	-1	$+1$	$+1$	$+1$
2	$+1$	-1	-1	-1
(c) $m=8$				
1	-1	-1	-1	-1
1	$+1$	$+1$	$+1$	$+1$
1	-1	-1	$+1$	$+1$
1	$+1$	$+1$	-1	-1
1	-1	$+1$	-1	$+1$
1	$+1$	-1	$+1$	-1
1	-1	$+1$	$+1$	-1
1	$+1$	-1	-1	$+1$
2	-1	-1	-1	$+1$
2	$+1$	$+1$	$+1$	-1
2	-1	-1	$+1$	-1
2	$+1$	$+1$	-1	$+1$
2	-1	$+1$	-1	-1
2	$+1$	-1	$+1$	$+1$
2	-1	$+1$	$+1$	$+1$
2	$+1$	-1	-1	-1

attributes. The size of the factorial design or the orthogonal fraction determines the number of profile sets for each block of the binary incomplete block design used to determine a pattern of the constant attributes. The design used for each combination of two constant attributes in the conjoint design in Table 1 is the full 2^2 factorial design. This implies that $2^2=4$ profile sets have the same pair of constant attributes with levels that form a full 2^2 factorial design.

If $k_c > 2$, minimum-size orthogonal fractions of the full 2^{k_c} factorial design can be used to fix the levels of the constant attributes. These fractions are the smallest possible ones that allow the estimation of k_c main effects. The design generators needed to construct the minimum-size orthogonal fractions for a specific number of constant attributes, k_c, can be found in Appendix 4A of Wu and Hamada (2000). An orthogonal fraction has a size of 2^{k_c-p}, where 2^{-p} refers to the fraction of the full 2^{k_c} factorial design. The full 2^{k_c-p} factorial design is used to determine the levels of the first k_c-p constant attributes. The levels of the remaining p constant attributes are specified by the generators.

The last problem to solve before the design construction is complete is the determination of the levels of the k_v non-constant attributes. For these attributes, we again use the orthogonal subdesigns given in the tables in Section 4.1. In the conjoint design of Table 1, the two possible subdesigns of Table 4a are implemented. Because that design is based on a binary incomplete block design with three blocks, three orthogonal subdesigns are required for the non-constant attributes. In Table 1, the first subdesign from Table 4a is used twice for the non-constant attributes (namely in profile sets 1–4 and in profile sets 9–12) while the other one is used once (namely in profile sets 5–8). The four blocks in these subdesigns match the four profile sets from the full 2^2 factorial design used for the levels of the constant attributes.

5. Information content of the designs

In this section, we examine five optimal variance-balanced designs that have $k=4$ attributes. We consider designs with $k_c=1$ constant attribute and $k_v=3$ non-constant attributes and designs with $k_c=2$ constant attributes and $k_v=2$ non-constant attributes. The designs with $k_c=1$ constant attribute have profile sets of sizes $m=2$, 4 and 8, respectively, while the designs with $k_c=2$ constant attributes have profile sets of sizes $m=2$ and 4. For all these designs, we computed the amount of information on the intercept and the part-worths, as measured by the diagonal elements of the information matrix in (8). The results appear in Table 5.

We compare the information content of each of the five conjoint designs in Table 5 with given values for n, b and m with that of an orthogonally blocked design involving $k=4$ two-level factors and n factor combinations arranged in b blocks of size m. Under the random respondent effects model (2), an orthogonally blocked design is a design that satisfies

$$\frac{1}{m}\mathbf{X}_i'\mathbf{1}_m = \frac{1}{n}\mathbf{X}'\mathbf{1}_n, \quad i=1,2,\ldots,b, \tag{14}$$

which means that the average row in each $\mathbf{X}_i$ is the same for all blocks and equal to the average row of the total model matrix $\mathbf{X}$ (Goos, 2002). For a two-level orthogonally blocked design and a main-effects model, the average rows in (14) equal $[1 \ \mathbf{0}_k']'$. We use the orthogonally blocked designs as benchmark because they are optimal and variance-balanced in the absence of constant attributes. They can be constructed using the generators in Appendix 3A and Appendix 4B of Wu and Hamada (2000). The designs also possess a diagonal information matrix in which the diagonal elements corresponding to the part-worths are all equal to n.

As the primary focus of researchers is on the estimates for the part-worths $\beta_1,\ldots,\beta_k$, we compare the values in the last column of Table 5 with the corresponding values for n, the information content of the benchmark designs. It is clear that the amount of information, n, on each part-worth obtained from a benchmark design, i.e. an orthogonally blocked design of size n, is larger than that obtained from any of the designs with constant attributes of size n. This is visualized in Fig. 1, where we plotted the percentage information losses incurred by using the five designs in Table 5 instead of the orthogonally blocked designs as a function of the degree of correlation ρ.

A close look at the percentages in Fig. 1 reveals that, for $m=2$ or 4 and a given value of ρ, the information losses for each part-worth of the conjoint designs with $k_c=2$ constant attributes are exactly twice the information losses of the conjoint designs with $k_c=1$ constant attribute. We verified that in general, if the number of attributes, k, is fixed in a series of

Table 5
Amount of information on the intercept and the part-worths obtained from the optimal variance-balanced conjoint designs, as measured by the diagonal elements of the information matrix, for $k_c=1$ and $k_v=3$ and for $k_c=k_v=2$.

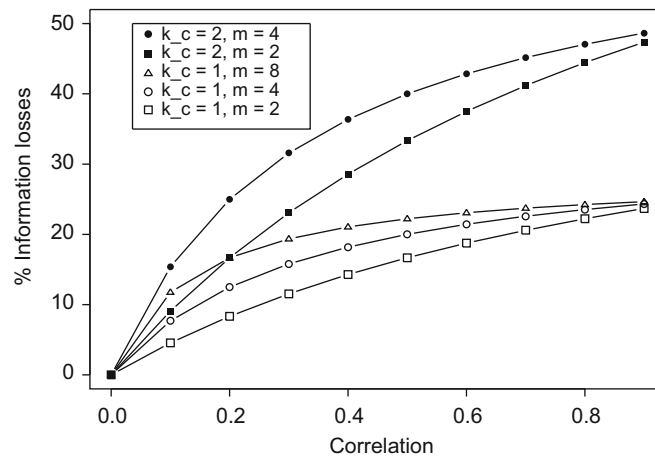| Design					Model parameter	
k_c	k_v	n	b	m	Intercept β_0	Part-worths $\beta_1,\ldots,\beta_k$
1	3	32	16	2	$32(1-\rho)/(1+\rho)$	$8(1-\rho)/(1+\rho)+24$
		32	8	4	$32(1-\rho)/(1+3\rho)$	$8(1-\rho)/(1+3\rho)+24$
		64	8	8	$64(1-\rho)/(1+7\rho)$	$16(1-\rho)/(1+7\rho)+48$
2	2	16	8	2	$16(1-\rho)/(1+\rho)$	$8(1-\rho)/(1+\rho)+8$
		32	8	4	$32(1-\rho)/(1+3\rho)$	$16(1-\rho)/(1+3\rho)+16$

Fig. 1. Percentage information losses per part-worth that result from using $k_c = 1$ and 2 constant attributes for degrees of correlation between 0 and 0.9.

optimal two-level variance-balanced conjoint designs, the information losses are proportional to the number of constant attributes, k_c, used. It is interesting to note that if $\rho = 1$, the percentage information losses per part-worth is equal to k_c/k.

Fig. 1 also shows that, for given values of k_c and ρ, the information loss for each part-worth increase with m. As the profile sets of the conjoint designs get larger or the number of respondents drops for a given n, the number of times the levels of the constant attributes vary over the design decreases which results in greater information losses from constant attributes. This can also be seen from the fact that $c_8 < c_4 < c_2 < 1$. It implies that for a given number of design profiles, n, sets with $m=2$ profiles are more efficient than sets with $m=4$ profiles, which in their turn, are more efficient than sets with $m=8$ profiles.

A final observation from Fig. 1 is that, given m and k_c, the information losses for each part-worth increase with ρ. The more heterogeneous respondents are, the more information one loses by keeping the levels of one or more attributes constant. Note that if $\rho = 0$, respondents are assumed to be homogeneous so that the grouping of the profiles in sets and the number of constant attributes, k_c, does not matter anymore.

6. Conclusion

In this paper, we presented a simple approach to construct optimal two-level conjoint designs that is especially useful when the number of attributes is large. To reduce the cognitive burden on the respondents in such cases, we suggest holding the levels of one or more attributes constant in each of the profile sets. The designs have sets of 2, 4 or 8 profiles and, for given values of n, b, m and k_c, they are optimal with respect to the $\mathcal{D}$-, $\mathcal{A}$-, $\mathcal{G}$- and $\mathcal{V}$-optimality criteria for estimating main-effects models.

The optimal conjoint designs have a diagonal information matrix and maximize the information content of the experiment. In addition, the conjoint designs are variance-balanced meaning that they yield an equal amount of information on each of the part-worths. Also, the conjoint designs do not depend on the extent to which respondents are heterogeneous, as expressed by the degree of correlation. This makes them very practical to use. However, compared with optimal conjoint designs without constant attributes, optimal conjoint designs with constant attributes lead to information losses that are proportional to the number of constant attributes.

If more than one constant attribute is desirable, our design construction method uses binary incomplete block designs with equal treatment replications to provide patterns of constant attributes that allow for variance balance. To ensure optimality, we choose the levels of the k_c constant attributes and the levels of the k_v non-constant attributes such that they are orthogonal for each combination of constant attributes. We first determine the levels of the k_c constant attributes by incorporating the full 2^{k_c} factorial design or an orthogonal fraction of it. We then set the levels of the k_v non-constant attributes by using a prespecified orthogonal subdesign for the k_v non-constant attributes.

A catalog of designs obtained using our approach can be found in Kessels et al. (2006). There are generally three ways to obtain larger conjoint designs. A first option is to choose binary incomplete block designs with larger numbers of blocks. A second possibility is to employ larger orthogonal designs to fix the levels of the k_c constant attributes. A third option is to replicate an entire design.

Finally, more work is needed to extend our method for constructing optimal two-level variance-balanced conjoint designs with sets of 3, 5, 6 or 7 profiles. Also the production of conjoint designs using models with main effects plus interactions might be considered.

Acknowledgements

Appendix

In this appendix we show that a design with a diagonal information matrix is better in terms of $\mathcal{D}$-, $\mathcal{A}$-, $\mathcal{G}$- and $\mathcal{V}$-optimality than a design with a non-diagonal information matrix containing the same diagonal elements as the diagonal information matrix. The $\mathcal{D}$- and $\mathcal{A}$-optimality criteria are both concerned with a precise estimation of the parameters in the model. A $\mathcal{D}$-optimal design minimizes the determinant of the variance–covariance matrix of the parameter estimates whereas an $\mathcal{A}$-optimal design minimizes the trace of the variance–covariance matrix. The $\mathcal{G}$- and $\mathcal{V}$-optimality criteria are concerned with making precise response predictions. The $\mathcal{G}$-optimality criterion seeks designs that minimize the maximum prediction variance over the region of interest, whereas the $\mathcal{V}$-optimality criterion seeks designs that minimize the average prediction variance over the region of interest.

Theorem. *A design with a diagonal information matrix* $\mathbf{M}^*$ *is better in terms of $\mathcal{D}$-, $\mathcal{A}$-, $\mathcal{G}$- and $\mathcal{V}$-optimality than any design with a non-diagonal information matrix* $\mathbf{M}$ *that has the same diagonal elements as* $\mathbf{M}^*$.

Proof. The proof is by induction. Consider the 2×2 information matrices

$$\mathbf{M}_2^* = \begin{bmatrix} a & 0 \\ 0 & c \end{bmatrix} \quad \text{and} \quad \mathbf{M}_2 = \begin{bmatrix} a & b \\ b & c \end{bmatrix},$$

and the $(p+1) \times (p+1)$ information matrices

$$\mathbf{M}_{p+1}^* = \begin{bmatrix} \mathbf{M}_p^* & \mathbf{0} \\ \mathbf{0} & c \end{bmatrix} \quad \text{and} \quad \mathbf{M}_{p+1} = \begin{bmatrix} \mathbf{M}_p & \mathbf{b} \\ \mathbf{b}' & c \end{bmatrix},$$

where $b \neq 0$, $\mathbf{b} \neq \mathbf{0}_p$ and $\mathbf{M}_p^*$ is a p-dimensional matrix with the same diagonal elements as $\mathbf{M}_p$.

The design corresponding to the information matrix $\mathbf{M}_2^*$ is better in terms of $\mathcal{D}$-optimality than that corresponding to $\mathbf{M}_2$ because $|\mathbf{M}_2^*| = ac > |\mathbf{M}_2| = ac - b^2$. Furthermore $|\mathbf{M}_{p+1}^*| = c|\mathbf{M}_p^*|$ and $|\mathbf{M}_{p+1}| = (c - \mathbf{b}'\mathbf{M}_p^{-1}\mathbf{b})|\mathbf{M}_p|$. Assuming that $|\mathbf{M}_p^*| > |\mathbf{M}_p|$ and noting that $c - \mathbf{b}'\mathbf{M}_p^{-1}\mathbf{b} < c$ because $\mathbf{M}_p^{-1}$ is positive definite, we can conclude that $|\mathbf{M}_{p+1}^*| > |\mathbf{M}_{p+1}|$ which proves the theorem with respect to $\mathcal{D}$-optimality.

The design corresponding to $\mathbf{M}_2^*$ is better than that corresponding to $\mathbf{M}_2$ in terms of $\mathcal{A}$-optimality because $\text{tr}(\mathbf{M}_2^{*-1}) = (a+c)/ac < \text{tr}(\mathbf{M}_2^{-1}) = (a+c)/(ac-b^2)$. Now, $\text{tr}(\mathbf{M}_{p+1}^{*-1}) = \text{tr}(\mathbf{M}_p^{*-1}) + 1/c$ and

$$\text{tr}(\mathbf{M}_{p+1}^{-1}) = \text{tr}\left[\mathbf{M}_p^{-1} + \frac{\mathbf{M}_p^{-1}\mathbf{b}\mathbf{b}'\mathbf{M}_p^{-1}}{c - \mathbf{b}'\mathbf{M}_p^{-1}\mathbf{b}}\right] + \frac{1}{c - \mathbf{b}'\mathbf{M}_p^{-1}\mathbf{b}}.$$

Assuming that $\text{tr}(\mathbf{M}_p^{*-1}) < \text{tr}(\mathbf{M}_p^{-1})$ and noting that

$$\text{tr}\left[\mathbf{M}_p^{-1} + \frac{\mathbf{M}_p^{-1}\mathbf{b}\mathbf{b}'\mathbf{M}_p^{-1}}{c - \mathbf{b}'\mathbf{M}_p^{-1}\mathbf{b}}\right] > \text{tr}(\mathbf{M}_p^{-1})$$

because $\mathbf{M}_p^{-1}\mathbf{b}\mathbf{b}'\mathbf{M}_p^{-1}$ is non-negative definite and that $c - \mathbf{b}'\mathbf{M}_p^{-1}\mathbf{b} < c$, we can conclude that $\text{tr}(\mathbf{M}_{p+1}^{*-1}) < \text{tr}(\mathbf{M}_{p+1}^{-1})$. The design with information matrix $\mathbf{M}_{p+1}^*$ is therefore better in terms of $\mathcal{A}$-optimality than the design corresponding to $\mathbf{M}_{p+1}$.

For main-effects models with or without interactions, the maximum prediction variance is attained at atleast one of the vertices of the design region. For a diagonal information matrix $\mathbf{M}_{p+1}^*$, $\mathbf{M}_{p+1}^{*-1}$ is also diagonal and the maximum prediction variance is equal to $\text{tr}(\mathbf{M}_{p+1}^{*-1})$. That maximum prediction variance is attained at each vertex of the design region. For a non-diagonal information matrix $\mathbf{M}_{p+1}$, the inverse $\mathbf{M}_{p+1}^{-1}$ is also non-diagonal, and, as shown above, it has a larger trace than $\mathbf{M}_{p+1}^{*-1}$. In addition, the maximum prediction variance corresponding to a non-diagonal $\mathbf{M}_{p+1}^{-1}$ is always larger than its trace. As a result, the maximum prediction variance corresponding to a non-diagonal information matrix is larger than that corresponding to a diagonal information matrix. This proves the theorem in terms of $\mathcal{G}$-optimality.

To complete the proof of the theorem, we now focus on the $\mathcal{V}$-optimality criterion. If we denote the design region by $\chi = [-1; +1]^m$, then a $\mathcal{V}$-optimal design is defined as the design that minimizes

$$\int_{\mathbf{x} \in \chi} \mathbf{f}'(\mathbf{x})\mathbf{M}^{-1}\mathbf{f}(\mathbf{x}) \, d\mathbf{x} = \int_{\mathbf{x} \in \chi} \text{tr}(\mathbf{f}'(\mathbf{x})\mathbf{M}^{-1}\mathbf{f}(\mathbf{x})) \, d\mathbf{x}$$

$$= \text{tr}\left(\mathbf{M}^{-1} \int_{\mathbf{x} \in \chi} \mathbf{f}(\mathbf{x})\mathbf{f}'(\mathbf{x}) \, d\mathbf{x}\right)$$

$$= \text{tr}(\mathbf{M}^{-1}\mathbf{D}),$$

where $\mathbf{D} = \int_{\mathbf{x} \in \chi} \mathbf{f}(\mathbf{x})\mathbf{f}'(\mathbf{x}) \, d\mathbf{x}$ is a diagonal matrix with positive diagonal elements $d_1, \dots, d_p$, the exact values of which depend on the number of experimental variables and the precise nature of the model. If we denote the diagonal elements of $\mathbf{M}^{-1}$

by m_i, the design criterion becomes

$$\int_{\mathbf{x}\in\chi} \mathbf{f}'(\mathbf{x})\mathbf{M}^{-1}\mathbf{f}(\mathbf{x})\,d\mathbf{x} = \sum_{i=1}^{p} m_i d_i.$$

Firstly, consider the situation where $p=2$. If a design's information matrix is the diagonal matrix $\mathbf{M}_2^*$ then its $\mathcal{V}$-criterion equals $v_2^*=d_1/a+d_2/c=(cd_1+ad_2)/(ac)$. If the information matrix of a design is given by $\mathbf{M}_2$, then its $\mathcal{V}$-criterion value equals $v_2=(cd_1+ad_2)/(ac-b^2)$ which is larger than v_2^*. Assume now that $\mathrm{tr}(\mathbf{M}_p^{*-1}\mathbf{D}) < \mathrm{tr}(\mathbf{M}_p^{-1}\mathbf{D})$ or, equivalently, that $v_p^* < v_p$. The $\mathcal{V}$-criterion value of a design with diagonal information matrix $\mathbf{M}_{p+1}^*$ equals $v_{p+1}^*=\mathrm{tr}(\mathbf{M}_p^{*-1}\mathbf{D})+d_{p+1}/c$. The $\mathcal{V}$-criterion value of a design with non-diagonal information matrix $\mathbf{M}_{p+1}$ equals

$$v_{p+1} = \mathrm{tr}(\mathbf{M}_p^{-1}\mathbf{D}) + \frac{\mathrm{tr}(\mathbf{M}_p^{-1}\mathbf{b}\mathbf{b}'\mathbf{M}_p^{-1}\mathbf{D})}{c-\mathbf{b}'\mathbf{M}_p^{-1}\mathbf{b}} + \frac{d_{p+1}}{c-\mathbf{b}'\mathbf{M}_p^{-1}\mathbf{b}}.$$

Since $c-\mathbf{b}'\mathbf{M}_p^{-1}\mathbf{b} < c$ and $\mathrm{tr}(\mathbf{M}_p^{-1}\mathbf{b}\mathbf{b}'\mathbf{M}_p^{-1}\mathbf{D})$ is positive we have that $v_{p+1} > v_{p+1}^*$. As a result, a design that corresponds to a diagonal information matrix is better in terms of $\mathcal{V}$-optimality than a design that corresponds to a non-diagonal information matrix with the same diagonal elements. □

A corollary of the theorem is that a design that maximizes the diagonal elements of the information matrix and that, at the same time, has a diagonal information matrix is $\mathcal{D}$-, $\mathcal{A}$-, $\mathcal{G}$- and $\mathcal{V}$-optimal.

References

Abel, R.J.R., Greig, M., 2006. BIBDs with small block size. In: Colbourn, C.J., Dinitz, J.H. (Eds.), The CRC Handbook of Combinatorial Designs. CRC Press, Boca Raton, FL, pp. 71–79 (Chapter I.2).

Brazier, J., Roberts, J., Deverill, M., 2002. The estimation of a preference-based measure of health from the SF-36. Journal of Health Economics 21, 271–292.

Clatworthy, W.H., 1973. Tables of Two-Associate-Class Partially Balanced Designs. Applied Mathematics Series 63. National Bureau of Standards (U.S.).

Cochran, W.G., Cox, G.M., 1957. Experimental Designs. Wiley, New York.

Cox, D.R., 1958. Planning of Experiments. Wiley, New York.

Danaher, P.J., 1997. Using conjoint analysis to determine the relative importance of service attributes measured in customer satisfaction surveys. Journal of Retailing 73, 235–260.

Goos, P., 2002. The Optimal Design of Blocked and Split-Plot Experiments. Springer, New York.

Goos, P., 2006. Optimal versus orthogonal and equivalent-estimation design of blocked and split-plot experiments. Statistica Neerlandica 60, 361–378.

Goos, P., Vandebroek, M., 2001. Optimal split-plot designs. Journal of Quality Technology 33, 436–450.

Goos, P., Vandebroek, M., 2003. D-optimal split-plot designs with given numbers and sizes of whole plots. Technometrics 45, 235–245.

Goos, P., Vandebroek, M., 2004. Outperforming completely randomized designs. Journal of Quality Technology 36, 12–26.

Grasshoff, U., Grossmann, H., Holling, H., Schwabe, R., 2003. Optimal paired comparison designs for first-order interactions. Statistics 37, 373–386.

Grasshoff, U., Grossmann, H., Holling, H., Schwabe, R., 2004. Optimal designs for main effects in linear paired comparison models. Journal of Statistical Planning and Inference 126, 361–376.

Green, P.E., 1974. On the design of choice experiments involving multifactor alternatives. Journal of Consumer Research 1, 61–68.

Grossmann, H., Holling, H., Brocke, M., Grasshoff, U., Schwabe, R., 2005. On the empirical relevance of optimal designs for the measurement of preferences. In: Berger, M., Wong, W.-K. (Eds.), Applied Optimal Designs. Wiley, New York, pp. 45–65.

Grossmann, H., Holling, H., Schwabe, R., 2002. Advances in optimum experimental design for conjoint analysis and discrete choice models. In: Franses, P.H., Montgomery, A.L. (Eds.), Advances in Econometrics, Econometric Models in Marketing, vol. 16. JAI Press, Amsterdam, pp. 93–117.

Kessels, R., Goos, P., Vandebroek, M., 2008. Optimal designs for conjoint experiments. Computational Statistics and Data Analysis 52, 2369–2387.

Kessels, R., Goos, P., Vandebroek, M., 2006. Optimal two-level conjoint designs for large numbers of attributes. Technical Report 0615, Faculty of Business and Economics, Katholieke Universiteit Leuven, 39pp.

Mathon, R., Rosa, A., 2006. 2-(v, k, λ) designs of small order. In: Colbourn, C.J., Dinitz, J.H. (Eds.), The CRC Handbook of Combinatorial Designs. CRC Press, Boca Raton, FL, pp. 25–58 (Chapter I.1).

Pullman, M.E., Moore, W.L., Wardell, D.G., 2002. A comparison of quality function deployment and conjoint analysis in new product design. Journal of Product Innovation Management 19, 354–364.

Schwabe, R., Grasshoff, U., Grossmann, H., Holling, H., 2003. Optimal 2^K paired comparison designs for partial profiles. In: Stulajter, F., Wimmer, G. (Eds.), PROBASTAT2002, Proceedings of the Fourth International Conference on Mathematical Statistics, Smolenice 2002. Tatra Mountains Mathematical Publications, vol. 26, pp. 79–86.

Shah, K.R., Sinha, B.K., 1989. Theory of Optimal Designs. Springer-Verlag, New York.

Sinha, K., Website of partially balanced incomplete block designs: ⟨http://sites.google.com/site/kishoresinha/partiallybalanceddesigns⟩.

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

Street, D.J., 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.

Street, D.J., Street, A.P., 1996. Partially balanced incomplete block designs. In: Colbourn, C.J., Dinitz, J.H. (Eds.), The CRC Handbook of Combinatorial Designs. CRC Press, Boca Raton, FL, pp. 419–423 (Chapter IV.35).

Wu, C.F.J., Hamada, M., 2000. Experiments: Planning, Analysis, and Parameter Design Optimization. Wiley, New York.