Journal of Choice Modelling, 4(3), 2011, 52-74
www.jocm.org.uk

Bayesian optimal designs for discrete choice experiments with partial profiles

Roselinde Kessels[1,*] Bradley Jones[2,1,†] Peter Goos[1,3,‡]

[1] Faculty of Applied Economics and StatUa Center for Statistics
Universiteit Antwerpen, Prinsstraat 13, 2000 Antwerpen, Belgium

[2] SAS Institute Inc., SAS Campus Drive, Cary, NC 27513, USA

[3] Erasmus School of Economics, Erasmus Universiteit Rotterdam
Postbus 1738, 3000 DR Rotterdam, The Netherlands

Received 2 May 2011, revised version received 7 September 2011, accepted 20 November 2011

Abstract

In a discrete choice experiment, each respondent chooses the best product or service sequentially from many groups or choice sets of alternative goods. The alternatives are described by levels of a set of predefined attributes and are also referred to as profiles. Respondents often find it difficult to trade off prospective goods when every attribute of the offering changes in each comparison. Especially in studies involving many attributes, respondents get overloaded by the complexity of the choice task. To overcome respondent fatigue, it is better to simplify the choice tasks by holding the levels of some of the attributes constant in every choice set. The resulting designs are called partial profile designs. In this paper, we construct D-optimal partial profile designs for estimating main-effects models. We use a Bayesian design algorithm that integrates the D-optimality criterion over a prior distribution of likely parameter values. To determine the constant attributes in each choice set, we generalize the approach that makes use of balanced incomplete block designs. Our algorithm is very flexible because it produces partial profile designs of any choice set size and allows for attributes with any number of levels and any number of constant attributes. We provide an illustration in which we make recommendations that balance the loss of statistical information and the burden imposed on the respondents.
Keywords: discrete choice experiments; Bayesian D-optimal design; partial profiles; lexicographic choice behavior; attribute balance; coordinate-exchange algorithm

*T: +32 (0)3 265 40 95, F: +32 (0)3 220 48 17, roselinde.kessels@ua.ac.be
†T: +1 919 531 4161, F: +1 919 677 4444, bradley.jones@jmp.com
‡T: +32 (0)3 220 40 59, F: +32 (0)3 220 48 17, peter.goos@ua.ac.be

ISSN 1755-5345

1 Introduction

Discrete choice experiments (DCEs) are widely used to study people's preferences for certain attributes of products or services in different applied fields such as marketing, transport, health economics and environmental economics. They are also called stated choice or conjoint choice experiments. Typically, DCEs involve respondents choosing among hypothetical (occasionally real) alternatives presented in choice sets where the alternatives are described by levels of a set of predefined attributes. The alternatives in the choice sets are also named profiles. The design of a DCE comprises a select number of choice sets administered to a group of respondents.

Respondents' choices are analyzed with discrete choice models in a random utility framework to estimate the preference parameters attached to each attribute level. The assumption implicit in these models is that respondents are willing to make compensatory decisions. This means that unattractive levels of an attribute can be compensated for by attractive levels of another attribute. If the assumption of compensatory decision making is violated, then this will have a detrimental effect on the validity of the estimated preference values.

Many practitioners use DCEs to study the relative importance of a large number of attributes (see, e.g., Sculpher et al. 2004; Witt et al. 2009). Constructing the design of a DCE involving many attributes is, however, not straightforward because respondents are limited in their cognitive ability to process the information presented in the DCE (Mazzotta and Opaluch 1995). Several studies exploring the complexity and cognitive burden associated with DCEs provide evidence that a large number of attributes has a detrimental effect on the ability to choose, contributing to an increased error variance. For example, Arentze et al. (2003) found that, when increasing the number of attributes from three to five in a transport mode choice study, the error variance increased substantially and the parameter estimates, corrected for the error variance differences, changed as well. Caussade et al. (2005) examined the use of 3, 4, 5 and 6 attributes through different designs for a route choice experiment, in which they also varied the number of alternatives in a choice set, the number of choice sets, the number of attribute levels and the range for those levels. They concluded that the number of attributes had the largest influence on the error variance out of all design dimensions.

According to Green (1974) and Schwabe et al. (2003), respondents get overwhelmed if they have to choose from a set of profiles that vary on more than four attributes. Restricting the number of attributes to a maximum of four bounds the increase in error variance, though we believe this cut-off number is rather arbitrary and should be re-evaluated for different DCEs. Nevertheless, it has been shown that, as respondents attempt to process many attributes, they often shortcut the compensatory process by making choices that are based on the levels of just one attribute or a small subset of the attributes (see, e.g., Hensher 2006; Hensher and Rose 2009; Scarpa et al. 2009). These attributes are typically the most important ones. They dominate the decision making process while the other attributes are ignored. The corresponding decision rule respondents apply then is non-compensatory.

To accurately measure respondents' trade-offs, it makes sense to simplify the comparison by holding the levels of some of the attributes constant in every choice set. These constant attributes need not be the same in each choice set. They can be ignored in the choice task so that the remaining attributes whose levels are varied make up the resulting choice set. The profiles in such a choice set are called partial profiles, and the number of attributes that are allowed to vary in the partial profiles is called the profile strength. The advantage of using partial profiles is that the compensatory discrete choice models remain valid because they help prevent respondents resorting to non-compensatory decision rules, such as lexicographic decision rules. Also, in the presence of a dominant attribute, we still obtain information about trade-offs made between the remaining attributes when partial profiles are used. The downside of partial profiles is that, in theory, they provide less information on the parameter values compared to full profiles that allow all attributes to vary in each choice set (Kessels et al. 2010).

It is interesting to note that the construction methods for full profile designs proposed in the choice design literature range from forcing the levels of each attribute to differ as much as possible between the profiles in each choice set to allowing for constant attributes, or complete attribute level overlap, in the choice sets. The most well-known full profile designs tying in with the former design approach are the designs by Street and Burgess. As pointed out by Rose and Bliemer (2009), the Street and Burgess designs may promote lexicographic choice behavior. Because the designs typically show all levels of each attribute in each choice set, a particularly dominant attribute level may govern every choice.

The contrasting approach to constructing full profile designs allows for attribute level overlap, or, in the extreme, complete attribute level overlap if an attribute has a constant level in a choice set. Attribute level overlap or constant attributes may be necessary in full profile designs to prevent the creation of choice sets with a dominant profile. Well-known design methods that allow for attribute level overlap are the locally and Bayesian optimal design methods that make use of prior information in the design construction. These methods are described later in this section. Kessels et al. (2011a,b) provide a detailed discussion comparing different full profile designs in terms of their numbers of constant attributes. Strictly speaking, if constant attributes are present in full profile designs, then the corresponding choice sets also have partial profiles. However, the term "partial profiles" is generally reserved to the systematic creation of a prespecified number of constant attributes in each choice set. As a result, systematically enforcing the use of constant attributes in every choice set leads to partial profile designs.

Thus far, partial profile designs have only been constructed for main-effects models under the convenient but unrealistic assumption that people have no preference for any of the profiles (Grasshoff et al. 2003, 2004; Grossmann et al. 2006, 2009). These so-called utility-neutral partial profile designs are created using orthogonal designs based on linear design principles. To improve the quality of the designs, innovative construction methods have been developed based on optimal design theory for linear models (see, e.g., Atkinson et al. 2007). Nevertheless, the most recent utility-neutral optimal designs created by Grossmann et al. (2009) only allow for the construction of DCEs with two profiles per choice set and two groups of attributes where the number of levels of the attributes is fixed in every group. There are also limitations on the allowable number of constant attributes. This approach is therefore not applicable in a wide variety of practical problems.

However, a more fundamental problem with the use of utility-neutral optimal designs is that they do not match the discrete choice models. This is because discrete choice models are nonlinear in the parameters, implying that the quality of the design of a DCE depends on the unknown parameters (Atkinson and Haines 1996). One justification for the use of the utility-neutral design approach is that the nonlinear design problem can be transformed into a linear one by assuming zero prior parameter values. A key feature of the utility-neutral optimal designs is thus that they are optimal for one specific set of parameter values. Therefore, utility-neutral optimal designs belong to the class of locally optimal designs which are constructed using prior point estimates for the parameters (Huber and Zwerina 1996).

To provide a more sound solution to the nonlinear design problem, we construct partial profile designs using the Bayesian design methodology. The Bayesian design approach, introduced in the choice design literature by Sándor and Wedel (2001), assumes a prior distribution of likely parameter values and optimizes the design over that distribution. In this way, it accounts for the uncertainty about the proposed parameters into the problem formulation. This leads to a better reflection of reality than in the utility-neutral design approach. Many researchers have implemented the Bayesian design approach to construct full profile designs for DCEs (see, e.g., Sándor and Wedel 2001, 2002; Scarpa and Rose 2008; Kessels et al. 2008, 2009, 2011a; Bliemer and Rose 2010). To create partial profile designs, we built on the work by Kessels et al. (2011a) who construct Bayesian full profile designs using the Bayesian $\mathcal{D}$-optimality criterion for the multinomial logit (MNL) model (McFadden 1974). These designs maximize the expected logarithm of the determinant of the information matrix of the maximum likelihood parameter estimators in the MNL model (Gotwalt et al. 2009). They are also referred to as $\mathcal{D}_B$-optimal designs. The models we consider are main-effects models only.

To generate $\mathcal{D}_B$-optimal partial profile designs, we use a two-stage design algorithm. In the first stage, we determine the constant attributes in each choice set, while in the second stage, we determine the levels of the non-constant attributes. To determine the constant attributes in each choice set, we maximize the $\mathcal{D}$-optimality criterion for a very specific fixed block effects model. This solution provides a generalization to the approach by Green (1974) who suggests using balanced incomplete block designs (BIBDs) to select the constant attributes. Our two-stage design algorithm is very flexible because it produces partial profile designs of any choice set size and allows for attributes with any number of levels and any number of constant attributes.

The outline of the remainder of the paper is as follows. Section 2 reviews the multinomial logit model and the $\mathcal{D}_B$-optimality criterion used to construct partial profile designs. In Section 3, we give an overview of the features of the Bayesian full profile design algorithm and present the Bayesian partial profile design algorithm. In Section 4, we adapt the algorithms for generating utility-neutral designs with full and partial profiles. Section 5 provides an illustration in which we construct and compare a series of $\mathcal{D}_B$-optimal partial profile designs. We study their performance relative to full profile testing and use utility-neutral designs as benchmarks. Section 6 concludes the paper and highlights some further research possibilities.

2 The multinomial logit framework

The multinomial logit (MNL) model (McFadden 1974) relies on random utility theory which describes the utility that respondent n ($n = 1, ..., N$) attaches to profile j ($j = 1, ..., J$) in choice set s ($s = 1, ..., S$) as the sum of a systematic and a stochastic component:

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

In the systematic component $\mathbf{x}'_{njs}\boldsymbol{\beta}$, $\mathbf{x}_{njs}$ is a $k \times 1$ vector containing the attribute levels of profile j in choice set s for respondent n. The vector $\boldsymbol{\beta}$ is a $k \times 1$ vector of parameter values representing the main effects of the attribute levels on the utility. The stochastic component ε_{njs} is the error term, which is assumed i.i.d. Gumbel distributed. The cumulative distribution function of an individual error term is

$$F(\varepsilon_{njs}) = \exp(-\exp(-\mu\varepsilon_{njs})), \tag{2}$$

where $\mu > 0$ is a scalar equal to $\pi/\sqrt{6}\sigma$ with σ the standard deviation of ε_{njs} (Ben-Akiva and Lerman 1985). The MNL probability that respondent n chooses profile j in choice set s is the closed-form expression

$$p_{njs} = \frac{\exp\left(\mathbf{x}'_{njs}\mu\boldsymbol{\beta}\right)}{\sum_{t=1}^{J}\exp\left(\mathbf{x}'_{nts}\mu\boldsymbol{\beta}\right)}. \tag{3}$$

The embedded scalar constant, μ, in the MNL model is the scale factor for a particular data set. It is inversely related to the size of the stochastic component and indicates what the quality of the data is. The smaller the value for μ, the larger the error variance σ^2, and vice versa. If μ approaches zero, then σ^2 approaches infinity, indicating a completely stochastic choice process with equal MNL probabilities. Conversely, if μ approaches infinity, then σ^2 approaches zero, indicating a deterministic choice process leaving no doubt about the preferred choice.

However, estimation of the MNL model (3) does not allow to separately identify both $\boldsymbol{\beta}$ and μ (Ben-Akiva and Lerman 1985; Swait and Louviere 1993; Louviere et al. 2002). Only the product $\mu\boldsymbol{\beta}$ can be estimated. Because of this identification problem, it is standard practice to normalize μ to 1. This normalization leads to the MNL model with choice probability

$$p_{njs} = \frac{\exp\left(\mathbf{x}'_{njs}\boldsymbol{\beta}\right)}{\sum_{t=1}^{J}\exp\left(\mathbf{x}'_{nts}\boldsymbol{\beta}\right)}, \tag{4}$$

where $\boldsymbol{\beta}$ is estimated using maximum likelihood. Because the same parameter vector $\boldsymbol{\beta}$ is attached to every respondent, it is assumed in this model that people's preferences for the attribute levels are homogeneous across the population.

The construction of an optimal design $\mathbf{X}_n = [\mathbf{x}'_{njs}]_{j=1,\ldots,J;s=1,\ldots,S}$ for respondent n for estimating $\boldsymbol{\beta}$ in the MNL model (4) is based on the Fisher information matrix

$$\mathbf{M}\left(\mathbf{X}_n, \boldsymbol{\beta}\right) = \sum_{s=1}^{S} \mathbf{X}'_{ns}\left(\mathbf{P}_{ns} - \mathbf{p}_{ns}\mathbf{p}'_{ns}\right)\mathbf{X}_{ns}, \tag{5}$$

with $\mathbf{X}_{ns} = [\mathbf{x}'_{njs}]_{j=1,\ldots,J}$ the submatrix of $\mathbf{X}_n$ corresponding to choice set s, $\mathbf{p}_{ns} = [p_{n1s}, ..., p_{nJs}]'$ and $\mathbf{P}_{ns} = \mathrm{diag}\,[p_{n1s}, ..., p_{nJs}]$. Huber and Zwerina (1996), Sándor and Wedel (2001), Kessels et al. (2006, 2009, 2011a), Scarpa and Rose (2008), among others, implemented different design criteria or functions of the information matrix (5) for constructing optimal discrete choice designs. This task is far from trivial since the information on $\boldsymbol{\beta}$ depends on the unknown values of $\boldsymbol{\beta}$ through the probabilities p_{njs} so that parameter values are required before it is possible to construct optimal designs. To deal with this dependency on $\boldsymbol{\beta}$, one can use a single prior guess, $\boldsymbol{\beta}_P$, in a locally optimal design approach. The most popular local design criterion is the $\mathcal{D}_P$-optimality criterion, which we define as

$$\mathcal{D}_P = \log\left|\mathbf{M}\left(\mathbf{X}_n, \boldsymbol{\beta}_P\right)\right|. \tag{6}$$

The design that maximizes the $\mathcal{D}_P$-criterion is the locally $\mathcal{D}$-optimal or $\mathcal{D}_P$-optimal design for the MNL model (4).

However, because of the uncertainty on $\boldsymbol{\beta}$, locally optimal designs are only efficient for parameters $\boldsymbol{\beta}$ close to $\boldsymbol{\beta}_P$. A more robust design solution is a Bayesian strategy that averages the design criterion over a prior distribution of likely parameter values, $\pi\left(\boldsymbol{\beta}\right)$. Often, this distribution is the multivariate normal distribution, $\mathcal{N}(\boldsymbol{\beta}|\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$, with prior mean $\boldsymbol{\beta}_0$ and prior variance-covariance matrix $\boldsymbol{\Sigma}_0$. As opposed to locally optimal designs, Bayesian optimal designs perform well for a broad range of parameters $\boldsymbol{\beta}$ (Kessels et al. 2011a,b; Rose 2011). To generate them, we use the $\mathcal{D}_B$-optimality criterion, which we define as

$$\mathcal{D}_B = \int_{\mathcal{R}^k} \log\left|\mathbf{M}\left(\mathbf{X}_n, \boldsymbol{\beta}\right)\right| \pi(\boldsymbol{\beta})\mathrm{d}\boldsymbol{\beta}. \tag{7}$$

The design that maximizes the $\mathcal{D}_B$-criterion is the Bayesian $\mathcal{D}$-optimal or $\mathcal{D}_B$-optimal design for the MNL model (4). Kessels et al. (2011a) were the first to use this definition of the $\mathcal{D}_B$-criterion to generate $\mathcal{D}_B$-optimal designs. It differs from the $\mathcal{D}_B$-optimality criterion used in most of the literature on optimal choice design (see, e.g., Sándor and Wedel 2001; Kessels et al. 2006, 2008, 2009; Bliemer and Rose 2010) because of the logarithmic transformation of the determinant. This transformation ensures that in a Bayesian information theoretic sense, the design that maximizes the $\mathcal{D}_B$-optimality criterion (7) also maximizes the expected Shannon information (Chaloner and Verdinelli 1995; Atkinson et al. 2007). A practical advantage of the logarithmic transformation is that it makes the $\mathcal{D}_B$-criterion less sensitive to parameter vectors resulting in very small or very large determinant values.

To compare the statistical efficiency of a $\mathcal{D}_B$-optimal partial profile design to the $\mathcal{D}_B$-optimal full profile design, we compute the relative $\mathcal{D}_B$-efficiency. We define the $\mathcal{D}_P$-efficiency of a design $\mathbf{X}_n$ relative to a design $\mathbf{X}_n^*$ as

$$Eff_P\left(\mathbf{X}_n, \mathbf{X}_n^*\right) = \exp\left(\frac{\mathcal{D}_P(\mathbf{X}_n) - \mathcal{D}_P(\mathbf{X}_n^*)}{k}\right). \tag{8}$$

By analogy, the $\mathcal{D}_B$-efficiency of a design $\mathbf{X}_n$ relative to a design $\mathbf{X}_n^*$ is

$$Eff_B\left(\mathbf{X}_n, \mathbf{X}_n^*\right) = \exp\left(\frac{\mathcal{D}_B(\mathbf{X}_n) - \mathcal{D}_B(\mathbf{X}_n^*)}{k}\right). \tag{9}$$

Also Holling and Schwabe (2011) proposed using Equation (9) to compute the relative $\mathcal{D}_B$-efficiency of two designs.

Rather than adopting a Bayesian design approach, some researchers have transformed the design problem for the MNL model (4), which is in essence nonlinear, into a linear one by creating locally optimal designs assuming $\boldsymbol{\beta} = \mathbf{0}_k$, where $\mathbf{0}_k$ is a k-dimensional vector of zeroes (see, e.g., Street and Burgess 2007; Grossmann et al. 2009). The assumption that $\boldsymbol{\beta} = \mathbf{0}_k$ causes the probabilities p_{njs} of all J profiles in choice set s to be equal to $1/J$, which corresponds to a situation in which respondents have no preference for any of the profiles in the choice set. Therefore, these designs are referred to as utility-neutral designs. Kessels et al. (2011a) showed that, up to a proportionality constant, the information matrix of a utility-neutral design equals the information matrix of a block design, which is given by

$$\mathbf{M}\left(\mathbf{X}_n\right) = \mathbf{X}_n'\mathbf{X}_n - \sum_{s=1}^{S} J^{-1}\left(\mathbf{X}_{ns}'\mathbf{1}_J\right)\left(\mathbf{1}_J'\mathbf{X}_{ns}\right), \tag{10}$$

where $\mathbf{1}_J$ is a J-dimensional vector of ones.

Therefore, to construct $\mathcal{D}$-optimal utility-neutral designs, we maximize the linear $\mathcal{D}$-optimality criterion,

$$\mathcal{D} = \left|\mathbf{X}_n'\mathbf{X}_n - \sum_{s=1}^{S} J^{-1}\left(\mathbf{X}_{ns}'\mathbf{1}_J\right)\left(\mathbf{1}_J'\mathbf{X}_{ns}\right)\right|. \tag{11}$$

3 Bayesian design algorithms

In this section, we describe the Bayesian design algorithms for generating $\mathcal{D}_B$-optimal full and partial profile designs. In Section 3.1, we review the features of the algorithm to construct $\mathcal{D}_B$-optimal full profile designs. We discuss this algorithm because it provides the basis for the algorithmic construction of $\mathcal{D}_B$-optimal partial profile designs, which we explain in Section 3.2.

3.1 Algorithm for generating $\mathcal{D}_B$-optimal full profile designs

The algorithm for generating $\mathcal{D}_B$-optimal full profile designs is an improvement to the alternating sample algorithm for generating Bayesian full profile designs developed by Kessels et al. (2009). Characteristic of the alternating sample algorithm is that it alternates between two different samples of prior parameters to approximate the k-dimensional integral related to a multivariate normal prior $\pi(\boldsymbol{\beta}) = \mathcal{N}(\boldsymbol{\beta}|\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$ in the definition of the $\mathcal{D}_B$-optimality criterion (7). The alternating sample algorithm uses a small systematic sample of prior parameters to generate design improvements in a random start, but still checks the designs produced by each random start using a large Monte Carlo sample. The approach works well because the small systematic sample agrees almost completely with the large Monte Carlo sample on design improvements in a random start.

The use of two different samples of prior parameters drastically speeded up the Bayesian design generation compared to using a large Monte Carlo sample only, as was done in Sándor and Wedel (2001) and Kessels et al. (2006). However, we increased the speed and precision of the algorithm even further by adopting the fast and accurate quadrature scheme presented by Gotwalt et al. (2009) and Gotwalt (2010) as integration method to compute the $\mathcal{D}_B$-optimality criterion (7). The quadrature scheme is based on the work of Monahan and Genz (1997) and combines two efficient integration rules, one for radial integration and one for integration over a sphere. The combined integration method is therefore called the spherical-radial transformation method. Bliemer et al. (2008) and Yu et al. (2010) showed that quadrature methods often outperform any other methods for Bayesian objective function evaluation.

Apart from using a more efficient sampling scheme from the multivariate normal distribution to compute the $\mathcal{D}_B$-optimality criterion (7), the Bayesian design algorithm works similarly to the alternating sample algorithm described in Kessels et al. (2009). It uses Meyer and Nachtsheim's (1995) coordinate-exchange algorithm that changes one coordinate or attribute level of a profile at a time. For each attribute level in the design, the coordinate-exchange algorithm tries all possible levels and chooses the level corresponding to the largest $\mathcal{D}_B$-criterion value. The algorithm runs several times through the design and restarts for a given number of randomly generated starting designs.

As opposed to the alternating sample algorithm, which has been developed for generating $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal full profile designs, we use the

Bayesian design algorithm to generate $\mathcal{D}_B$-optimal full profile designs only. These designs are most popular because they are easier to compute and do not only guarantee precise parameter estimates, but also precise predictions (Kessels et al. 2006). The $\mathcal{D}_B$-optimal full profile design algorithm has been implemented in the statistical software package JMP, which we use to generate the designs. Also Kessels et al. (2011a) generated a series of $\mathcal{D}_B$-optimal full profile designs in this way.

3.2 Algorithm for generating $\mathcal{D}_B$-optimal partial profile designs

This section discusses the adaptation of the Bayesian full profile design algorithm described in Section 3.1 so that it allows for the creation of $\mathcal{D}_B$-optimal partial profile designs. Given there are t attributes under study, the $\mathcal{D}_B$-optimal partial profile design algorithm requires the number of constant attributes, t_c, as an input. The remaining $t_v = t - t_c$ attributes are the non-constant attributes.

The Bayesian partial profile design algorithm for t_c constant attributes is characterized by the following two stages:

Stage 1: Determining the constant attributes in each choice set,

Stage 2: Determining the levels of the non-constant attributes.

Note that for main-effects models, the levels of the t_c constant attributes can be chosen randomly because they have no effect on the information acquired from the experiment, as expressed by the information matrix (5). We now describe each of the two stages in more detail.

3.2.1 Stage 1: Determining the constant attributes in each choice set

To determine the constant attributes in each choice set in Stage 1, we attempt to balance the number of times an attribute is held constant in the choice design. Also, if $t_c > 1$, we attempt to balance the number of times an attribute is held constant with another attribute. We refer to the first type of balance as first-order balance and to the second type of balance as second-order balance or pairwise balance. We refer to the entire approach as attribute balance.

As Green (1974) proposed, we can use balanced incomplete block designs (BIBDs) to determine patterns of t_c constant attributes that perfectly satisfy attribute balance. By definition, a BIBD describes how to arrange the levels of a single qualitative factor, called treatments, in groups or blocks of a certain size. Each treatment thereby occurs an equal number of times in the entire design and the number of times two different treatments occur together in a block is the same for all pairs of treatments.

Table 1 shows a $\mathcal{D}_B$-optimal partial profile design containing 15 choice sets of size 2 for $t = 6$ attributes where $t_c = 2$ attributes are constant in each choice set. The first three attributes have two levels each, the fourth and fifth attribute have three levels each and the last attribute has five levels. The $t_c = 2$ constant attributes are marked in gray and can assume any possible attribute level. As such, the choice design has a profile strength of four. We selected the $t_c = 2$ constant attributes according to a BIBD that has 15 blocks with $t_c = 2$ of the $t = 6$ attributes as treatments in each block. The BIBD ensures first-order balance because each attribute is constant in five choice sets and second-order balance because each pair of attributes is constant in exactly one choice set.

A problem with using BIBDs as a way to determine the t_c constant attributes in choice designs is that they only exist for certain numbers of observations, treatments and block sizes. To circumvent this problem, we propose a more general $\mathcal{D}$-optimal design approach that satisfies attribute balance to the largest possible extent in every design situation. This $\mathcal{D}$-optimal design approach is based on the fixed block effects model for data from a BIBD,

$$Y_{uv} = \mathbf{q}'_{uv}\boldsymbol{\alpha} + \mathbf{z}'_u\boldsymbol{\gamma} + \epsilon_{uv}, \tag{12}$$

where Y_{uv} denotes the response of the vth treatment ($v = 1, ..., t_c$) in the uth block ($u = 1, ..., S$). Next, $\mathbf{q}_{uv}$ is a $t \times 1$ vector that has a one as first element followed by $t-1$ effects-type coded elements representing the vth treatment in the uth block. The vector $\boldsymbol{\alpha} = [\alpha_0, \alpha_1, \ldots, \alpha_{t-1}]'$ contains the corresponding t parameters with α_0 the intercept and $\alpha_1, \ldots, \alpha_{t-1}$ the treatment effects. Then, $\mathbf{z}_u$ is a $(S-1)\times 1$ vector containing effects-type coded elements for the uth block and $\boldsymbol{\gamma}$ is the corresponding $(S-1)\times 1$ vector containing the fixed block effects. Finally, the term ϵ_{uv} is the random error associated with the vth treatment in the uth block, which is assumed to be independently distributed with zero mean and variance σ^2_ϵ.

In matrix notation, the fixed block effects model (12) becomes

$$\mathbf{Y} = \mathbf{Q}\boldsymbol{\alpha} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\epsilon}, \tag{13}$$

where $\mathbf{Y}$ is a vector of $r = St_c$ responses, $\mathbf{Q} = [\mathbf{q}'_{uv}]_{u=1,\ldots,S;v=1,\ldots,t_c}$ is the $r \times t$ design matrix having a vector of ones in the first column and treatment codings in the remaining columns, $\mathbf{Z} = [\mathbf{1}_{t_c} \otimes z'_u]_{u=1,\ldots,S}$ is the $r \times (S-1)$ design matrix of associated block codings and $\boldsymbol{\epsilon}$ is a random error vector. We denote by $\mathbf{B}$ the combined design matrix $[\mathbf{Q}\ \mathbf{Z}]$ which has dimension $r \times (t + S - 1)$.

The fixed block effects model (13) is linear in the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$, implying that the information matrix of the design $\mathbf{B}$ is independent of their values. The information matrix is given by

$$\mathbf{B}'\mathbf{B} = \begin{bmatrix} \mathbf{Q}'\mathbf{Q} & \mathbf{Q}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{Q} & \mathbf{Z}'\mathbf{Z} \end{bmatrix}. \tag{14}$$

By maximizing $|\mathbf{B}'\mathbf{B}|$, we obtain the $\mathcal{D}$-optimal design for the fixed block effects model (13). This design will be a BIBD if a BIBD exists for the specified values of S, t_c and t. Otherwise, if no BIBD exists, the $\mathcal{D}$-optimal design will be as close as possible to a BIBD.

Table 1: $\mathcal{D}_B$-optimal partial profile design involving 15 choice sets with 2 profiles and three 2-level attributes, two 3-level attributes and one 5-level attribute, where $t_c = 2$ of the attributes are constant.

Choice set	Attributes 1	2	3	4	5	6
1	*	*	2	1	1	5
1	*	*	1	3	3	3
2	*	1	*	2	3	5
2	*	2	*	3	1	1
3	*	2	1	*	2	1
3	*	1	2	*	1	2
4	*	2	2	2	*	2
4	*	1	1	3	*	5
5	*	2	1	2	1	*
5	*	1	2	3	3	*
6	1	*	*	2	3	1
6	2	*	*	1	2	3
7	2	*	1	*	1	3
7	1	*	2	*	2	4
8	1	*	2	2	*	3
8	2	*	1	3	*	2
9	2	*	1	1	3	*
9	1	*	2	3	1	*
10	2	1	*	*	1	4
10	1	2	*	*	2	2
11	1	2	*	3	*	5
11	2	1	*	1	*	4
12	2	1	*	2	2	*
12	1	2	*	1	3	*
13	1	2	1	*	*	4
13	2	1	2	*	*	1
14	2	2	2	*	1	*
14	1	1	1	*	2	*
15	1	1	2	1	*	*
15	2	2	1	2	*	*

Based on the information matrix (14), we can compute $|\mathbf{B}'\mathbf{B}|$ by applying Theorem 13.3.8 of Harville (1997) to find that

$$\begin{vmatrix} \mathbf{Q}'\mathbf{Q} & \mathbf{Q}'\mathbf{Z} \\ \mathbf{Z}'\mathbf{Q} & \mathbf{Z}'\mathbf{Z} \end{vmatrix} = |\mathbf{Z}'\mathbf{Z}| \left|\mathbf{Q}'\mathbf{Q} - \mathbf{Q}'\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\mathbf{Q}\right| . \tag{15}$$

Because the matrix $\mathbf{Z}$ only depends on S and t_c, it is constant. Also, $|\mathbf{Z}'\mathbf{Z}|$ and $\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'$ are constant so that we only need to compute them once. If we denote $\mathbf{Z}\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'$ by $\mathbf{H}$ and the r-dimensional identity matrix by $\mathbf{I}_r$, then in our algorithm, we maximize $|\mathbf{B}'\mathbf{B}|$ by maximizing

$$|\mathbf{Q}'(\mathbf{I}_r - \mathbf{H})\mathbf{Q}| . \tag{16}$$

This approach leads to computational time savings that increase with the number of blocks, S.

3.2.2 Stage 2: Determining the levels of the non-constant attributes

To determine the levels of the t_v non-constant attributes in Stage 2, we apply the same methods as in the Bayesian full profile design algorithm described in Section 3.1. We generate changes in the levels of the non-constant attributes using the coordinate-exchange algorithm and we evaluate these changes using the $\mathcal{D}_B$-optimality criterion (7). To compute the $\mathcal{D}_B$-criterion value of a design for the multivariate normal prior $\pi(\boldsymbol{\beta}) = \mathcal{N}(\boldsymbol{\beta}|\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$, we rely on the spherical-radial transformation method presented by Gotwalt et al. (2009) and Gotwalt (2010).

In the $\mathcal{D}_B$-optimal partial profile design for $t_c = 2$ constant attributes shown in Table 1, we observe that the levels of all $t_v = 4$ non-constant attributes vary in the choice sets. However, in some other instances, we observed that the levels of one or more of the t_v attributes we allow to vary do not change in certain choice sets. In that case, we see more than the specified minimum number of attributes remaining constant in such choice sets. This result is in line with Sándor and Wedel (2002) and Kessels et al. (2006, 2011a,b) who also found the presence of constant attributes, or complete attribute level overlap, in some choice sets of full profile designs. Using constant attributes in certain choice sets of full profile designs helps avoiding uninformative choice sets with very high and very low choice probabilities.

The number of constant attributes selected by the $\mathcal{D}_B$-optimality criterion (7) in addition to the t_c constant attributes depends primarily on the number of t_c constant attributes already present. Additionally, the specification of the prior mean $\boldsymbol{\beta}_0$ and prior variance-covariance matrix $\boldsymbol{\Sigma}_0$ in the multivariate normal distribution plays a role. Kessels et al. (2011a) showed that the larger the prior variances and the Euclidean distance $d(\boldsymbol{\beta}_0, \mathbf{0}_k)$ of the prior mean $\boldsymbol{\beta}_0$ to the zero vector $\mathbf{0}_k$, the more constant attributes the $\mathcal{D}_B$-optimality criterion selects in a full profile design. As a result, for partial profile designs, the $\mathcal{D}_B$-optimality criterion is more likely to select constant attributes to complement the

t_c constant attributes given large values for the prior variances and $d(\boldsymbol{\beta}_0, \mathbf{0}_k)$. In the illustration in Section 5, we describe the prior parameter distribution used to generate the partial profile design of Table 1 and discuss the example in more detail.

In many DCEs, it is, however, desirable that the number of constant attributes is the same in each choice set. In particular, in DCEs where the profiles are displayed as rows or columns of words or sentences and the constant attributes are dropped, the rows or columns do not all have the same length if the number of constant attributes is not the same in each choice set. This is aesthetically jarring. On the other hand, if the profiles are shown as images or real-life prototypes (e.g. images or prototypes of products), then the use of choice sets with different numbers of constant attributes is generally not a problem. In that case, the images or prototypes do not immediately reveal how many attributes are constant or non-constant.

For DCEs that require partial profile designs with a fixed number of t_c constant attributes, we modified the partial-profile coordinate-exchange algorithm so that it does not allow for additional constant attributes selected by the $\mathcal{D}_B$-optimality criterion. We thus restrict the levels of the non-constant attributes to vary in the choice sets. We refer to the entire two-stage algorithm as the restricted partial profile design algorithm and to the resulting designs as restricted partial profile designs. Accordingly, we can then also call the original version of the algorithm, without the restriction, the unrestricted partial profile design algorithm and the resulting designs unrestricted partial profile designs. The unrestricted partial profile design algorithm has been implemented in JMP. For the restricted partial profile design algorithm, we created our own script using JMP Scripting Language (JSL).

4 Utility-neutral design algorithms

As described in Section 2, utility-neutral designs are in essence block designs with blocks of size J. For that reason, we generate $\mathcal{D}$-optimal utility-neutral full profile designs by maximizing the linear $\mathcal{D}$-optimality criterion (11). The algorithm we use is the coordinate-exchange algorithm which is implemented in JMP. For main-effects models, the linear $\mathcal{D}$-optimality criterion varies the levels of all attributes. Moreover, the resulting $\mathcal{D}$-optimal utility-neutral full profile designs are level balanced within and over all choice sets given appropriate numbers of attribute levels, blocks and block sizes. For other design situations, the main-effects designs are level balanced to the largest possible extent.

To generate $\mathcal{D}$-optimal utility-neutral partial profile designs for t_c constant attributes, we apply a similar two-stage design scheme as described in Section 3.2 for generating Bayesian partial profile designs. In Stage 1, we select the constant attributes so as to achieve attribute balance. In Stage 2, we generate the levels of the non-constant attributes by maximizing the linear $\mathcal{D}$-optimality criterion (11) using the coordinate-exchange algorithm. Also here, the linear $\mathcal{D}$-optimality criterion varies the levels of all t_v attributes for main-effects models, so that the partial profile designs have exactly t_c attributes constant. Furthermore, these main-effects designs are level balanced in the non-constant attributes within and over all choice sets for appropriate numbers of attribute levels, blocks and block sizes. For other design dimensions, the designs are level balanced to the largest possible extent. The utility-neutral partial profile design algorithm has been implemented in JMP.

5 Illustration

This section provides an illustrative study in which we compare a series of $\mathcal{D}_B$-optimal partial profile designs to the corresponding $\mathcal{D}_B$-optimal full profile design. Other benchmark designs are the $\mathcal{D}$-optimal utility-neutral designs with full and partial profiles. All designs are main-effects designs. They consist of 15 choice sets with 2 profiles and three 2-level attributes, two 3-level attributes and one 5-level attribute. We assume a single design for all N respondents, so that $\mathbf{X}_1 = ... = \mathbf{X}_N$ for every design case. We examine the loss in $\mathcal{D}_B$-efficiency due to using partial profiles instead of full profiles. We describe the setup of the comparison study in Section 5.1 and discuss the optimal designs and their $\mathcal{D}_B$-efficiencies in Section 5.2.

5.1 Setup of the comparison study

In Section 3.2, we described our two-stage design procedure to construct $\mathcal{D}_B$-optimal partial profile designs using the $\mathcal{D}_B$-optimal partial profile design for $t_c = 2$ constant attributes shown in Table 1. In this comparison study, we discuss this design in more detail and present the unrestricted and restricted $\mathcal{D}_B$-optimal partial profile design for $t_c = 1$ constant attribute. We compare these partial profile designs with the $\mathcal{D}_B$-optimal full profile design in terms of $\mathcal{D}_B$-efficiency. We also study the $\mathcal{D}$-optimal utility-neutral design with full profiles, and the $\mathcal{D}$-optimal utility-neutral partial profile designs for $t_c = 1$ and $t_c = 2$ constant attributes.

We generated the $\mathcal{D}_B$-optimal and $\mathcal{D}$-optimal utility-neutral designs using the algorithms described in Sections 3 and 4, respectively. For each optimal design, we used 2000 random starts. We modeled the attribute levels using effects-type coding. If L_i denotes the number of levels for attribute i, $i = 1, ..., t$, then effects-type coding constrains the parameters associated with the L_i levels of attribute i to sum to zero. This requires that only the parameters attached to the first $L_i - 1$ levels of attribute i need to be estimated as the parameter attached to the last level L_i automatically results. Therefore, the vector $\boldsymbol{\beta}$ contains $k = 11$ unknown parameter values in our design situation.

For the construction of the Bayesian designs, we used the multivariate normal prior $\pi(\boldsymbol{\beta}) = \mathcal{N}(\boldsymbol{\beta}|\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$, with prior mean

$$\boldsymbol{\beta}_0 = [-0.5, -0.5, -0.5, -0.5, 0, -0.5, 0, -0.5, -0.25, 0, 0.25]' \tag{17}$$

and prior variance-covariance matrix

$$\boldsymbol{\Sigma}_0 = \begin{bmatrix} 0.16 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.16 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.16 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.16 & -0.08 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.08 & 0.16 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.16 & -0.08 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -0.08 & 0.16 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.16 & -0.04 & -0.04 & -0.04 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.04 & 0.16 & -0.04 & -0.04 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.04 & -0.04 & 0.16 & -0.04 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.04 & -0.04 & -0.04 & 0.16 \end{bmatrix}. \tag{18}$$

For the specification of the prior mean $\boldsymbol{\beta}_0$, we chose to equally space the parameter values between -0.5 and 0.5 for each attribute. The first three parameter values in $\boldsymbol{\beta}_0$ are the prior mean utilities associated with the first level of each of the three 2-level attributes. The next two sets of two parameter values reflect the prior mean utilities of the first and second level of each of the two 3-level attributes. Finally, the last four parameter values correspond to the prior mean utilities of the first four levels of the 5-level attribute. Due to the effects-type coding, a prior mean parameter value of 0.5 automatically results for the last level of each attribute.

By using the prior mean $\boldsymbol{\beta}_0$ in (17), we assume that all $t = 6$ attributes are equally important and that they are ordinal, where the levels of each attribute are ordered from least preferred to most preferred. Because there are $t = 6$ attributes involved in the study, which is reasonably large, we followed the recommendation by Kessels et al. (2008) to use small absolute prior mean parameter values. None of the parameter values is therefore larger than 0.5 in magnitude. As a result, the multinomial logit probabilities given $\boldsymbol{\beta}_0$ for the most and least desirable full profile forming a choice set of size two are 0.99753 and 0.00247, respectively, which are values that are not too extreme.

For the prior variance-covariance matrix $\boldsymbol{\Sigma}_0$, we specified $k = 11$ variances that are all equal to 0.16 and negative covariances between the $L_i - 1$ parameters of each attribute i. We computed these covariances using a correlation coefficient of $-1/(L_i - 1)$. As explained by Kessels et al. (2008), this ensures that the variances of all parameters corresponding to a given attribute are the same, meaning that the variance associated with the last implied level L_i of attribute i also equals 0.16.

5.2 Optimal designs and their $\mathcal{D}_B$-efficiencies

Table 2 shows the $\mathcal{D}_B$-optimal full profile design and the unrestricted and restricted $\mathcal{D}_B$-optimal partial profile design for $t_c = 1$ constant attribute. The unrestricted design has four choice sets (namely choice sets 3, 4, 6 and 12) with two constant attributes instead of one. The additional four constant attributes imposed by the $\mathcal{D}_B$-optimality criterion (7) involve the three two-level attributes only. In the partial profile designs, the attribute balance in the constant attributes is not perfect because the number of choice sets, $S = 15$, is not a multiple of the number of attributes, $t = 6$. Also, in the unrestricted design, the attribute balance is not perfect because of the four additional constant attributes.

Notice that in the $\mathcal{D}_B$-optimal full profile design, there are seven choice sets with one constant 2-level attribute. There are 6 attributes and 15 choice sets giving 90 possibilities for such level overlap. Here, we see about 8% of level overlap, which is rather small compared to the percentages from a series of design instances reported by Kessels et al. (2011a). This is due to the fairly small prior variances and covariances in the variance-covariance matrix $\mathbf{\Sigma}_0$ in (18) and the comparatively small Euclidean distance $d(\boldsymbol{\beta}_0, \mathbf{0}_{11})$ of the prior mean $\boldsymbol{\beta}_0$ in (17) to the zero vector $\mathbf{0}_{11}$, which equals 1.27. In the case of $t_c = 1$ constant attribute, we were therefore somewhat surprised to see some additional constant attributes in the unrestricted $\mathcal{D}_B$-optimal design. On the other hand, in the case of $t_c = 2$ constant attributes, the unrestricted partial profile design of Table 1 does not have more constant attributes than required. As a result, the generation of a restricted partial profile design for this situation is unnecessary.

Table 3 shows the $\mathcal{D}$-optimal utility-neutral full profile design and the $\mathcal{D}$-optimal utility-neutral partial profile designs for $t_c = 1$ and $t_c = 2$ constant attributes. Characteristic of the utility-neutral designs is that they are level balanced in the non-constant attributes within and over all choice sets. Similar to the Bayesian partial profile designs, the use of $S = 15$ choice sets and $t = 6$ attributes resulted in imperfect attribute balance for the utility-neutral design with $t_c = 1$ constant attribute and in perfect attribute balance for the utility-neutral design with $t_c = 2$ constant attributes. If there are one or more dominant attributes in the study, then the full profile design does not prevent the selection of profiles based on these dominant attributes, whereas the partial profile designs do to some extent.

To evaluate the statistical efficiency of the Bayesian and utility-neutral optimal designs, we compared them to the $\mathcal{D}_B$-optimal full profile design in terms of $\mathcal{D}_B$-efficiency. By doing so, we learn how much we lose in $\mathcal{D}_B$-efficiency by using constant attributes and/or a $\mathcal{D}$-optimal utility-neutral design. Table 4 shows the $\mathcal{D}_B$-optimal criterion values and $\mathcal{D}_B$-efficiencies of the designs. In the case of $t_c = 1$ constant attribute, the $\mathcal{D}_B$-optimal partial profile designs lose about 10% in $\mathcal{D}_B$-efficiency. Also, there is hardly any difference in performance between the unrestricted and restricted design. This means that in order to obtain the same amount of information as from full profile testing, we need about 10% more re-

Table 2: $\mathcal{D}_B$-optimal designs involving 15 choice sets with 2 profiles and three 2-level attributes, two 3-level attributes and one 5-level attribute.

	Full profile design						1 cst attribute partial profile designs											
							Unrestricted						Restricted					
Choice	Attributes						Attributes						Attributes					
set	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
1	1	2	*	2	3	1	*	1	2	2	2	5	*	1	1	2	3	4
1	2	1	*	3	1	5	*	2	1	3	1	4	*	2	2	1	1	2
2	*	1	2	3	2	2	*	1	2	3	1	3	*	2	1	3	1	1
2	*	2	1	2	1	5	*	2	1	2	3	4	*	1	2	1	2	3
3	*	1	2	2	3	2	*	*	2	2	1	2	1	*	1	3	3	2
3	*	2	1	3	2	1	*	*	1	1	3	5	2	*	2	1	1	1
4	2	1	1	1	3	3	*	*	2	1	1	1	1	*	2	1	3	2
4	1	2	2	3	1	2	*	*	1	3	2	2	2	*	1	3	2	3
5	1	1	*	2	2	5	1	*	2	1	2	1	2	*	1	1	1	4
5	2	2	*	1	3	1	2	*	1	2	1	5	1	*	2	2	2	1
6	1	2	1	3	3	2	1	*	*	3	3	3	1	2	*	2	3	3
6	2	1	2	2	2	4	2	*	*	1	2	2	2	1	*	3	2	4
7	1	2	2	1	2	5	1	2	*	2	2	1	1	1	*	3	3	5
7	2	1	1	3	1	3	2	1	*	1	3	2	2	2	*	2	1	4
8	2	1	1	2	3	5	1	2	*	3	1	5	1	2	*	2	1	5
8	1	2	2	1	2	3	2	1	*	2	3	3	2	1	*	1	3	1
9	1	2	*	1	1	4	1	2	2	*	3	2	1	2	1	*	2	4
9	2	1	*	3	2	1	2	1	1	*	2	4	2	1	2	*	1	5
10	1	1	2	1	3	5	1	1	2	*	3	4	1	1	2	*	3	4
10	2	2	1	2	1	4	2	2	1	*	2	3	2	2	1	*	2	2
11	2	*	1	1	3	4	2	1	2	*	2	4	2	1	1	*	1	3
11	1	*	2	2	1	1	1	2	1	*	3	2	1	2	2	*	2	4
12	2	1	2	1	1	1	2	1	*	3	*	1	1	2	1	1	*	5
12	1	2	1	3	2	3	1	2	*	1	*	3	2	1	2	2	*	2
13	2	2	1	2	2	2	2	2	1	3	*	1	2	1	1	2	*	5
13	1	1	2	3	1	4	1	1	2	2	*	4	1	2	2	3	*	3
14	1	1	1	3	3	4	1	1	1	3	2	*	1	1	2	2	2	*
14	2	2	2	1	1	3	2	2	2	1	1	*	2	2	1	3	3	*
15	2	*	1	1	2	2	1	2	2	3	2	*	1	1	2	3	1	*
15	1	*	2	2	3	3	2	1	1	2	1	*	2	2	1	1	3	*

Table 3: $\mathcal{D}$-optimal utility-neutral designs involving 15 choice sets with 2 profiles and three 2-level attributes, two 3-level attributes and one 5-level attribute.

	Full profile design						Partial profile designs											
							1 cst attribute						2 cst attributes					
Choice set	Attributes						Attributes						Attributes					
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
1	2	1	2	1	2	3	*	2	2	2	3	1	*	*	2	1	3	3
1	1	2	1	3	1	5	*	1	1	1	1	3	*	*	1	2	1	2
2	1	1	1	3	1	2	*	1	1	2	3	5	*	1	*	2	3	4
2	2	2	2	2	2	1	*	2	2	1	1	2	*	2	*	1	1	1
3	2	2	1	1	1	1	2	*	2	2	2	2	*	1	2	*	2	2
3	1	1	2	3	2	3	1	*	1	1	1	1	*	2	1	*	1	4
4	1	1	1	2	3	3	2	*	2	1	3	3	*	1	1	2	*	3
4	2	2	2	3	2	2	1	*	1	3	2	2	*	2	2	1	*	4
5	1	2	1	1	3	2	1	*	2	2	1	4	*	1	1	3	1	*
5	2	1	2	3	1	5	2	*	1	3	2	5	*	2	2	2	2	*
6	1	1	2	1	3	4	2	1	*	2	1	1	1	*	*	3	2	5
6	2	2	1	2	1	3	1	2	*	3	3	4	2	*	*	1	3	2
7	1	1	2	2	1	1	2	1	*	1	3	2	2	*	2	*	1	3
7	2	2	1	1	2	5	1	2	*	3	2	3	1	*	1	*	3	1
8	1	2	1	2	2	4	1	2	*	2	1	5	1	*	1	1	*	5
8	2	1	2	3	3	1	2	1	*	3	3	4	2	*	2	3	*	1
9	1	2	2	1	3	5	1	2	1	*	3	2	2	*	1	1	2	*
9	2	1	1	2	2	2	2	1	2	*	2	4	1	*	2	2	1	*
10	1	1	2	2	2	5	1	1	2	*	3	3	1	1	*	*	2	4
10	2	2	1	3	3	3	2	2	1	*	1	4	2	2	*	*	3	5
11	1	1	1	1	2	1	2	2	2	*	1	5	2	1	*	2	*	1
11	2	2	2	2	3	4	1	1	1	*	2	1	1	2	*	3	*	2
12	2	1	2	1	1	2	2	2	2	3	*	1	1	1	*	1	1	*
12	1	2	1	3	2	1	1	1	1	2	*	4	2	2	*	3	3	*
13	1	2	2	3	1	4	1	1	2	1	*	5	2	1	2	*	*	5
13	2	1	1	2	3	5	2	2	1	2	*	3	1	2	1	*	*	3
14	1	2	2	1	1	3	2	2	1	1	2	*	1	1	2	*	3	*
14	2	1	1	3	3	4	1	1	2	3	1	*	2	2	1	*	2	*
15	1	2	2	2	3	2	1	2	2	1	2	*	2	1	1	3	*	*
15	2	1	1	1	1	4	2	1	1	3	1	*	1	2	2	2	*	*

spondents (computed as $1/0.91-1$) if we keep one attribute constant. In the case of $t_c = 2$ constant attributes, the $\mathcal{D}_B$-optimal partial profile design loses about 20% in $\mathcal{D}_B$-efficiency, requiring 27% more respondents compared to full profile testing. This is twice the efficiency loss of the $\mathcal{D}_B$-optimal partial profile designs with $t_c = 1$ constant attribute. Also, this efficiency loss is approximately as large as the efficiency loss incurred by the $\mathcal{D}$-optimal utility-neutral full profile design.

Table 4: $\mathcal{D}_B$-optimal criterion values and $\mathcal{D}_B$-efficiencies of the optimal designs.

Optimal design			
Criterion	Profile	$\mathcal{D}_B$-value	$\mathcal{D}_B$-efficiency
$\mathcal{D}_B$-optimal	full	11.24378	100.00%
$\mathcal{D}_B$-optimal	1 cst partial	10.21495	91.07%
$\mathcal{D}_B$-optimal	1 cst partial restricted	10.20127	90.96%
$\mathcal{D}_B$-optimal	2 cst partial	8.64183	78.94%
$\mathcal{D}$-optimal utility-neutral	full	8.77159	79.87%
$\mathcal{D}$-optimal utility-neutral	1 cst partial	7.16088	68.99%
$\mathcal{D}$-optimal utility-neutral	2 cst partial	6.36122	64.15%

To evaluate the efficiency loss of 20% of the $\mathcal{D}$-optimal utility-neutral full profile design, we need to take into account the fact that the outperformance of Bayesian designs over utility-neutral designs depends on the prior parameter specification used. Kessels et al. (2011a) showed that when the prior variances are relatively small and the Euclidean distance $d(\boldsymbol{\beta}_0, \mathbf{0}_k)$ is relatively large, Bayesian designs perform much better than utility-neutral designs with respect to the variance of the parameter estimates. In our example, the prior variances all equal 0.16, which is relatively small, but the distance $d(\boldsymbol{\beta}_0, \mathbf{0}_{11})$ of 1.27 is not large. This means that the efficiency loss of 20% of the $\mathcal{D}$-optimal utility-neutral full profile design is not extreme, but rather average. Lastly, using $t_c = 1$ and $t_c = 2$ constant attributes in the utility-neutral design case yields efficiency losses of 31% and 36%, respectively, which makes these design options the worst to consider.

6 Summary and future research

We provided a flexible design algorithm for constructing $\mathcal{D}_B$-optimal main-effects designs for DCEs that reduce the complexity of the choices that a respondent must make by keeping one or more attributes constant within each choice set. We refer to these designs as $\mathcal{D}_B$-optimal partial profile designs. The benefits of the $\mathcal{D}_B$-optimal partial profile designs are twofold. First, by reducing the complexity of the choices, they have the potential to prevent respondents from resorting to non-compensatory approaches towards making their choices, which would violate the assumption of compensatory decision making in discrete choice models. Second, having attributes that remain constant within each choice set means that even if one attribute is dominant over all others, it will not appear in certain choice sets. Thus, the utility of the remaining attribute levels can be assessed.

Keeping certain attributes constant in each choice set also has a cost. It reduces the theoretical information content of each choice set compared to using full profiles. In our illustration, these efficiency losses were 10% to 20%, meaning that to obtain the same amount of information as from full profile testing, an investigator will need 10% to 27% more respondents. However, this theoretical drawback is outweighed by the potential of partial profile designs to prevent non-compensatory decision making. It would be interesting to verify these results by comparing designs with full and partial profiles in an empirical setting.

The partial profile design algorithm is flexible in that it can accommodate arbitrarily many attributes, each with any number of levels. Choice sets may have any number of profiles and though the number of choice sets must be adequate to fit the underlying model, there can be as many choice sets as desired. These may be divided into separate surveys so that each respondent is not overburdened by having to make too many choices. Our illustration had choice sets with two profiles only. Quantifying the efficiency of scenarios with more profiles per choice set would be a natural extension of this work.

Also, a number of more general extensions are worth investigating. First, the partial profile design algorithm considers main-effects models only. Examining the use of constant attributes when possible two-attribute interactions are present would be an interesting research topic. Second, the Bayesian design algorithm only allows for a multivariate normal prior distribution of the parameters. Extending the method to support other distributions would be a useful contribution. Finally, the algorithm only applies to the multinomial logit model assuming homogeneous preferences of the respondents. Exploring the use of more sophisticated nonlinear models that take into account respondent heterogeneity would be a challenging topic for future research. For example, the work by Bliemer and Rose (2010) and Yu et al. (2011) on full profile design construction methods for the panel mixed logit model could be extended to partial profiles.

References

Arentze, T., Borgers, A., Timmermans, H. and Del Mistro, R., 2003. Transport stated choice responses: effects of task complexity, presentation format and literacy, *Transportation Research E* **39**: 229–244.

Atkinson, A. C., Donev, A. N. and Tobias, R. D., 2007. *Optimum Experimental Designs, with SAS*, Oxford, U.K.: Clarendon Press.

Atkinson, A. C. and Haines, L. M., 1996. Designs for nonlinear and generalized linear models, in *Handbook of Statistics: Design and Analysis of Experiments*, Vol. 13, Ghosh, S. and Rao, C. R., eds. Amsterdam: Elsevier, 437–475.

Ben-Akiva, M. and Lerman, S., 1985. *Discrete Choice Analysis: Theory and Application* *to Travel Demand*, Cambridge MA: MIT Press.

Bliemer, M. C. J. and Rose, J. M., 2010. Construction of experimental designs for mixed logit models allowing for correlation across choice observations, *Transportation Research B* **44**: 720–734.

Bliemer, M. C. J., Rose, J. M. and Hess, S., 2008. Approximation of Bayesian efficiency in experimental choice designs, *Journal of Choice Modelling* **1**: 98–127.

Caussade, S., Ortúzar, J. de D., Rizzi, L. I. and Hensher, D. A., 2005. Assessing the influence of design dimensions on stated choice experiment estimates, *Transportation Research B* **39**: 621–640.

Chaloner, K. and Verdinelli, I., 1995. Bayesian experimental design: a review, *Statistical Science* **10**: 273–304.

Gotwalt, C. M., 2010. Addendum to "Fast computation of designs robust to parameter uncertainty for nonlinear settings", *Technometrics* **52**: 137.

Gotwalt, C. M., Jones, B. A. and Steinberg, D. M., 2009. Fast computation of designs robust to parameter uncertainty for nonlinear settings, *Technometrics* **51**: 88–105.

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

Grasshoff, U., Grossmann, H., Holling, H. and 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 multi-factor alternatives, *Journal of Consumer Research* **1**: 61–68.

Grossmann, H., Grasshoff, U. and Schwabe, R., 2009. Approximate and exact optimal designs for paired comparisons of partial profiles when there are two groups of factors, *Journal of Statistical Planning and Inference* **139**: 1171–1179.

Grossmann, H., Holling, H., Grasshoff, U. and Schwabe, R., 2006. Optimal designs for asymmetric linear paired comparisons with a profile strength constraint, *Metrika* **64**: 109–119.

Harville, D. A., 1997. *Matrix Algebra from a Statistician's Perspective*, New York: Springer.

Hensher, D. A., 2006. How do respondents process stated choice experiments? Attribute consideration under varying information load, *Journal of Applied Econometrics* **21**: 861–878.

Hensher, D. A. and Rose, J. M., 2009. Simplifying choice through attribute preservation or non-attendance: Implications for willingness to pay, *Transportation Research E* **45**: 583–590.

Holling, H. and Schwabe, R., 2011. Discussion on "The usefulness of Bayesian optimal designs for discrete choice experiments", *Applied Stochastic Models in Business and Industry* **27**: 189–192.

Huber, J. and Zwerina, K., 1996. The importance of utility balance in efficient choice designs, *Journal of Marketing Research* **33**: 307–317.

Kessels, R., Goos, P. and Vandebroek, M., 2006. A comparison of criteria to design efficient choice experiments, *Journal of Marketing Research* **43**: 409–419.

Kessels, R., Goos, P. and Vandebroek, M., 2010. Optimal two-level conjoint designs with constant attributes in the profile sets, *Journal of Statistical Planning and Inference* **140**: 3035–3046.

Kessels, R., Jones, B., Goos, P. and Vandebroek, M., 2008. Recommendations on the use of Bayesian optimal designs for choice experiments, *Quality and Reliability Engineering International* **24**: 737–744.

Kessels, R., Jones, B., Goos, P. and Vandebroek, M., 2009. An efficient algorithm for constructing Bayesian optimal choice designs, *Journal of Business and Economic Statistics* **27**: 279–291.

Kessels, R., Jones, B., Goos, P. and Vandebroek, M., 2011a. The usefulness of Bayesian optimal designs for discrete choice experiments, *Applied Stochastic Models in Business and Industry* **27**: 173–188.

Kessels, R., Jones, B., Goos, P. and Vandebroek, M., 2011b. Rejoinder: The usefulness of Bayesian optimal designs for discrete choice experiments, *Applied Stochastic Models in Business and Industry* **27**: 197–203.

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

Mazzotta, M. J. and Opaluch, J. J., 1995. Decision making when choices are complex: a test of Heiner's hypothesis, *Land Economics* **71**: 500–515.

McFadden, D., 1974. Conditional logit analysis of qualitative choice behavior, in Zarembka, P., *Frontiers in Econometrics*, New York: Academic Press, 105–142.

Meyer, R. K. and Nachtsheim, C. J., 1995. The coordinate-exchange algorithm for constructing exact optimal experimental designs, *Technometrics* **37**: 60–69.

Monahan, J. and Genz, A., 1997. Spherical-radial integration rules for Bayesian computation, *Journal of the American Statistical Association* **92**: 664–674.

Rose, 2011. Discussion on "The usefulness of Bayesian optimal designs for discrete choice experiments", *Applied Stochastic Models in Business and Industry* **27**: 193–196.

Rose, J. M. and Bliemer, M. C. J., 2009. Constructing efficient stated choice experimental designs, *Transport Reviews* **29**: 587–617.

Sándor, Z. and Wedel, M., 2001. Designing conjoint choice experiments using managers' prior beliefs, *Journal of Marketing Research* **38**: 430–444.

Sándor, Z. and Wedel, M., 2002. Profile construction in experimental choice designs for mixed logit models, *Marketing Science* **21**: 455–475.

Scarpa, R., Gilbride, T. J., Campbell, D. and Hensher, D. A., 2009. Modelling attribute non-attendance in choice experiments for rural landscape valuation, *European Review of Agricultural Economics* **36**: 151–174.

Scarpa, R. and Rose, J. M., 2008. Design efficiency for non-market valuation with choice modelling: how to measure it, what to report and why, *The Australian Journal of Agricultural and Resource Economics* **52**: 253–282.

Schwabe, R., Grasshoff, U., Grossmann, H. and Holling, H., 2003. Optimal 2^K paired comparison designs for partial profiles, in *PROBASTAT2002, Proceedings of the Fourth International Conference on Mathematical Statistics*, Smolenice 2002, Tatra Mountains Mathematical Publications, Vol. 26, Stulajter, F. and Wimmer, G., eds., 79–86.

Sculpher, M., Bryan, S., Fry, P., de Winter, P., Payne, H. and Emberton, M., 2004. Patients' preferences for the management of non-metastatic prostate cancer: discrete choice experiment, *British Medical Journal* **328**: 382–384.

Street, D. J. and Burgess, L., 2007. *The Construction of Optimal Stated Choice Experiments: Theory and Methods*, Hoboken, New Jersey: Wiley.

Swait, J. and Louviere, J. J., 1993. The role of the scale parameter in the estimation and comparison of multinomial logit models, *Journal of Marketing Research* **30**: 305–314.

Witt, J., Scott, A. and Osborne, R. H., 2009. Designing choice experiments with many attributes. An application to setting priorities for orthopaedic waiting lists, *Health Economics* **18**: 681–696.

Yu, J., Goos, P. and Vandebroek, M., 2010. Comparing different sampling schemes for approximating the integrals involved in the efficient design of stated choice experiments, *Transportation Research B* **44**: 1268–1289.

Yu, J., Goos, P. and Vandebroek, M., 2011. Individually adapted sequential Bayesian conjoint-choice designs in the presence of consumer heterogeneity, *International Journal of Research in Marketing* **28**: 378–388.