DEPARTMENT OF APPLIED ECONOMICS

RESEARCH REPORT

OPTIMAL DESIGNS FOR RATING-BASED CONJOINT EXPERIMENTS

ROSELINDE KESSELS • PETER GOOS • MARTINA VANDEBROEK

OR 0464

KATHOLIEKE UNIVERSITEIT LEUVEN

Optimal designs for rating-based conjoint experiments

Roselinde Kessels*
Peter Goos†
Martina Vandebroek*

Abstract

The scope of conjoint experiments on which we focus embraces those experiments in which each of the respondents receives a different set of profiles to rate. Carefully designing these experiments involves determining how many and which profiles each respondent has to rate and how many respondents are needed. To that end, the set of profiles offered to a respondent is viewed as a separate block in the design and a respondent effect is incorporated in the model, representing the fact that profile ratings from the same respondent are correlated. Optimal conjoint designs are then obtained by means of an adapted version of the algorithm of Goos and Vandebroek (2004). For various instances, we compute the optimal conjoint designs and provide some practical recommendations.

Keywords: conjoint analysis, optimal block design, rating-based conjoint experiments, $\mathcal{D}$-optimality

1 Introduction

In marketing, conjoint experiments have frequently been carried out to measure consumer preferences for the attributes of various products or services (Green et al. 2001). They have been conducted for issues of new product development, pricing, advertising, and other areas across many different industrial sectors around the world (Wittink and Cattin 1989, Wittink et al. 1994, Gustafsson et al. 2003). In a traditional conjoint experiment, respondents are presented with a set of alternatives or profiles that are defined as combinations of different attribute levels. Respondents are then requested to express their likelihood of purchase for each of the profiles by ranking them or by rating them on a point or monetary scale. Besides rating profiles on a scale, rating may also occur by means of directly asking reservation prices for the profiles. A reservation price for a good is the

*Department of Applied Economics, Faculty of Economics and Applied Economics, Katholieke Universiteit Leuven, Naamsestraat 69, B-3000 Leuven, Belgium.

†Department of Mathematics, Statistics & Actuarial Sciences, Faculty of Applied Economics, Universiteit Antwerpen, Prinsstraat 13, B-2000 Antwerpen, Belgium.

highest price a consumer is willing to pay for the good. Note that responses obtained from ranking are non-metric, whereas those obtained from rating are metric. The corresponding experiments are referred to as ranking-based and rating-based conjoint experiments.

Nowadays, rating-based conjoint experiments are most frequently employed (Wittink et al. 1994, Huber et al. 2003). This is mainly due to practical considerations such as the ease with which they can be used in telephone-interviews, web-based surveys, hybrid models and standard software packages. Therefore, in this paper, we devote all our attention to rating-based conjoint experiments.

In general, a conjoint experiment proceeds by submitting a small number of profiles to a limited number of respondents. In doing so, it aims at collecting as much information as possible on the utilities the respondents attach to the attributes of the product or service under investigation. To that end, an efficient experimental design needs to be developed. Most often, a different set of alternatives is given to each respondent as this enables the variation in the dependent variable to be better captured. However, the literature on conjoint experiments is silent about how to carefully design different sets of profiles given the number of experimental observations. For example, if 30 observations are available, then it is unclear whether it is statistically more efficient to submit ten alternatives to three respondents or three alternatives to ten respondents. Hence, we ask ourselves how many and which profiles each respondent has to rate and how many respondents are needed. Remark that we have to make sure that the conjoint designs constructed still remain practical. The number of profiles administered to a respondent should be kept small in order not to give in on response quality (Malhotra, 1986). On the other hand, if there are costs involved per respondent needed, the total costs may not become excessive. If statistically efficient conjoint designs turn out to be impractical, some compromise has to be made between statistical efficiency and practicality.

The approach we adopt to solve the conjoint problem is an algorithmic one and is based on the theory of optimal block designs, thoroughly dealt with in Goos and Vandebroek (2001a) and Goos (2002). Block designs are heavily used in industrial experimentation where observations are correlated. The model assumed is the linear random block effects model. This model incorporates a random effect representing the variation caused by the commonality on which the observations are grouped. In Goos (2002), optimal block designs are in depth compared to optimal completely randomized designs in terms of statistical efficiency. In contrast to block designs, observations in a completely randomized design are assumed to be uncorrelated and therefore, they are selected based on the linear model, without a block effect. Consequently, the observations are not grouped and can be randomly carried out. The comparison between the optimal block designs and the optimal completely randomized designs clearly demonstrated that in the case of correlated observations, optimal block designs are more efficient than optimal completely randomized designs.

Similar to block designs, we use the linear random block effects model to construct efficient conjoint designs. The motivation for this model in the conjoint setting is as follows. It is reasonable to assume that respondents are randomly selected from a population and that they are heterogeneous. Examples of respondent heterogeneity are the variations in terms of age, experience with the product or service under study, physical characteristics, cognitive abilities, etc. The consequence of this heterogeneity is that profile ratings from the same respondent are correlated. As a result, the random block effect in the model represents the effect of a respondent who has to rate a block of profiles. The efficient conjoint design for this model then consists of blocks of profiles that are each offered to a different respondent and the optimal number of respondents is derived from the number of blocks the design contains. We refer to the linear random block effects model in the conjoint setting as the linear random respondent effects model. Remark that we focus on main effects models only.

Other approaches to tackle respondent heterogeneity can be found in Liski et al. (2002) and Entholzner et al. (2004). For repeated measurement situations, they mathematically derived efficient designs using the linear random coefficient regression model, which allows for individual-specific regression parameters. Furthermore, to design and estimate conjoint experiments, Lenk et al. (1996) applied the hierarchical Bayes random effects model with subject-level covariates. Finally, in Cochran and Cox (1957), balanced incomplete block designs are recommended for preference rating as most of the design plans contain blocks with six or fewer units. Regrettably, these designs are only built on the levels of one qualitative experimental factor. These different levels are referred to as treatments. Together with balanced complete block designs, the designs belong to the class of balanced block designs which have the property that any pair of treatments appears together in a block equally often. Note that balanced complete block designs are not taken into consideration for preference rating as they generally contain blocks with more than six profiles. Typical for the balanced block designs is that they are universally optimal for the estimation of the treatment and the block effect. In other words, a balanced block design is optimal with respect to any generalized optimality criterion. Besides the fact that only one factor is taken up in the design, another disadvantage of balanced incomplete block designs is that they can only be used for specific numbers of observations, treatments and blocks. Consequently, for design situations in which no balanced incomplete block design is available, a.o. situations with more than one factor, optimal conjoint designs need to be computed.

The paper is organized as follows. First, we embark on the random respondent effects model in Section 2 and on the derivation of optimal conjoint designs in Section 3. Next, Section 4 presents the design construction algorithm and Section 5 discusses the computational results. Finally, Section 6 summarizes the paper.

2 The random respondent effects model

Suppose that a conjoint experiment consists of n profiles to be rated. In total, b respondents are appointed who each rate a different block or set of profiles. The profiles are correspondingly arranged in b blocks of sizes $m_1, ..., m_b$ with $n = \sum_{i=1}^{b} m_i$. If we assume that the respondents are random subjects from a population, then the utility u_{ij} for the jth profile provided by the ith respondent can be modelled as

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

In this model, $\mathbf{x}_{ij}$ is the $k \times 1$ vector having one as first element and the attribute levels describing profile j that is rated by respondent i as the remaining $k-1$ elements. $\boldsymbol{\beta} = [\beta_1, ..., \beta_k]'$ is the $k \times 1$ parameter vector with β_1 the intercept and $\beta_2, ..., \beta_k$ the weights associated with the attribute levels, hereby indicating the importance of the levels as viewed by the average respondent. γ_i represents the random effect of the ith respondent triggered by the heterogeneity of the respondents and ε_{ij} is the random error term. Note that the attributes are quantitative or discrete factors and that their levels are coded by means of the so-called effects-type coding which is often employed in the literature. If the number of levels for an attribute equals l, then the attribute is characterized by $l-1$ elements in $\mathbf{x}_{ij}$ and $\boldsymbol{\beta}$. To illustrate, in the case of a three-level attribute, the first level is coded as [1 0], the second level as [0 1] and the third level as [-1 -1]. For a two-level attribute, the codes are -1 for the first level and 1 for the second level. However, other types of coding may also be used as the choice of coding type does not affect the relative design efficiency in case the $\mathcal{D}$-optimality criterion is used (see Section 3).

In matrix notation, model (1) becomes

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

where $\mathbf{u}$ is a vector of n profile ratings, the vector $\boldsymbol{\beta}$ contains the k unknown fixed parameters, the vector $\boldsymbol{\gamma} = [\gamma_1, ..., \gamma_b]'$ contains the b random respondent effects and $\boldsymbol{\varepsilon}$ is a random error vector. The matrices $\mathbf{X}$ and $\mathbf{Z}$ have dimensions $n \times k$ and $n \times b$ respectively. $\mathbf{X}$ is given by

$$\mathbf{X} = [\mathbf{X}'_1, ..., \mathbf{X}'_b]', \tag{3}$$

where $\mathbf{X}_i = [\mathbf{x}_{i1}, ..., \mathbf{x}_{im_i}]'$ collects the profiles rated by respondent i and $\mathbf{Z}$ is defined as

$$\mathbf{Z} = \text{diag}[\mathbf{1}_{m_1}, ..., \mathbf{1}_{m_b}], \tag{4}$$

where $\mathbf{1}_{m_i}$ is a $m_i \times 1$ vector of ones. It is assumed that

$$\text{E}(\boldsymbol{\varepsilon}) = \mathbf{0}_n \text{ and } \text{Cov}(\boldsymbol{\varepsilon}) = \sigma^2_{\varepsilon}\mathbf{I}_n, \tag{5}$$

$$\text{E}(\boldsymbol{\gamma}) = \mathbf{0}_b \text{ and } \text{Cov}(\boldsymbol{\gamma}) = \sigma^2_{\gamma}\mathbf{I}_b, \tag{6}$$

$$\text{and } \text{Cov}(\boldsymbol{\gamma}, \boldsymbol{\varepsilon}) = \mathbf{0}_{b \times n}, \tag{7}$$

where σ^2_ε is the variability within respondents and σ^2_γ is the variability between respondents. We refer to this model as the random respondent effects model. Under these assumptions, the variance-covariance matrix of the observations $\mathbf{u}$ can be written as

$$\begin{aligned}
\mathbf{V} &= \text{Cov}(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}), \\
&= \text{Cov}(\mathbf{Z}\boldsymbol{\gamma}) + \text{Cov}(\boldsymbol{\varepsilon}), \\
&= \sigma^2_\gamma \mathbf{Z}\mathbf{Z}' + \sigma^2_\varepsilon \mathbf{I}_n, \\
&= \sigma^2_\gamma \text{diag}[\mathbf{1}_{m_1}\mathbf{1}'_{m_1}, ..., \mathbf{1}_{m_b}\mathbf{1}'_{m_b}] + \sigma^2_\varepsilon \text{diag}[\mathbf{I}_{m_1}, ..., \mathbf{I}_{m_b}], \\
&= \text{diag}[\sigma^2_\varepsilon \mathbf{I}_{m_1} + \sigma^2_\gamma \mathbf{1}_{m_1}\mathbf{1}'_{m_1}, ..., \sigma^2_\varepsilon \mathbf{I}_{m_b} + \sigma^2_\gamma \mathbf{1}_{m_b}\mathbf{1}'_{m_b}].
\end{aligned}$$

Substituting

$$\mathbf{V}_i = \sigma^2_\varepsilon \mathbf{I}_{m_i} + \sigma^2_\gamma \mathbf{1}_{m_i}\mathbf{1}'_{m_i}, \tag{8}$$

leads to the variance-covariance matrix

$$\mathbf{V} = \text{diag}[\mathbf{V}_1, ..., \mathbf{V}_b]. \tag{9}$$

It is interesting to observe that the block matrices $\mathbf{V}_i$ are compound symmetric: the main diagonals of these matrices contain the constant variances of the profile ratings, while the off-diagonal elements are constant covariances. For example, the variance-covariance matrix of five profile ratings u_{ij} obtained from two respondents, one of whom rated a block of two profiles and the other a block of three profiles, equals

$$\mathbf{V} = \begin{bmatrix}
\sigma^2_\varepsilon + \sigma^2_\gamma & \sigma^2_\gamma & 0 & 0 & 0 \\
\sigma^2_\gamma & \sigma^2_\varepsilon + \sigma^2_\gamma & 0 & 0 & 0 \\
0 & 0 & \sigma^2_\varepsilon + \sigma^2_\gamma & \sigma^2_\gamma & \sigma^2_\gamma \\
0 & 0 & \sigma^2_\gamma & \sigma^2_\varepsilon + \sigma^2_\gamma & \sigma^2_\gamma \\
0 & 0 & \sigma^2_\gamma & \sigma^2_\gamma & \sigma^2_\varepsilon + \sigma^2_\gamma
\end{bmatrix}.$$

The zero entries show that profile ratings from different respondents are assumed to be uncorrelated. The coefficient of correlation between two profile ratings from the same respondent is equal to

$$\rho = \frac{\sigma^2_\gamma}{\sigma^2_\varepsilon + \sigma^2_\gamma}. \tag{10}$$

This ratio $\rho \in [0, 1]$ measures the proportion of the total variability that is accounted for by the differences between respondents. If $\rho \to 0$, or equivalently $\sigma^2_\gamma \to 0$, then the profile ratings from the same respondent are no longer correlated. In this case, $\gamma_1 = ... = \gamma_b = 0$ and as a result, the random respondent effects model degenerates to the uncorrelated model

$$\mathbf{u} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}. \tag{11}$$

3 Analysis & design optimality

When the error terms as well as the respondent effects are normally distributed, the maximum likelihood estimator of the unknown fixed model parameters $\boldsymbol{\beta}$ in (1) and (2) is the generalized least squares (GLS) estimator. The GLS estimator is the best linear unbiased estimator (BLUE) and is given by

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{u}, \tag{12}$$

with variance-covariance matrix

$$\text{Cov}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}. \tag{13}$$

Sometimes, the variances components σ^2_γ and σ^2_ε are known from previous experimentation so that the estimator $\hat{\boldsymbol{\beta}}$ and its variance-covariance matrix can be immediately obtained. Most often, however, the variances σ^2_γ and σ^2_ε are unknown and therefore, (12) and (13) cannot be applied directly. Instead, the variance components σ^2_γ and σ^2_ε have to be estimated, for example via restricted maximum likelihood (REML) (Gilmour and Trinca 2000). The estimates $\hat{\sigma}^2_\gamma$ and $\hat{\sigma}^2_\varepsilon$ are then substituted in the GLS estimator (12), yielding the so-called feasible GLS estimator

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{u}, \tag{14}$$

where

$$\hat{\mathbf{V}} = \hat{\sigma}^2_\varepsilon\mathbf{I}_n + \hat{\sigma}^2_\gamma\mathbf{Z}\mathbf{Z}'.$$

In that case, the variance-covariance matrix (13) can be approximated by

$$\text{Cov}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1}. \tag{15}$$

The derivation of optimal designs for conjoint experiments is related to the variance-covariance matrix (13), and thus also to its inverse, the Fisher information matrix, on which we elaborate. The Fisher information matrix is denoted by

$$\mathcal{I}(\mathbf{X}) = \mathbf{X}'\mathbf{V}^{-1}\mathbf{X}. \tag{16}$$

In order to obtain a maximum of information on the parameters, we evaluate different design options by means of the $\mathcal{D}$-optimality criterion. The $\mathcal{D}$-optimality criterion is by far the most frequently used optimality criterion and it is a direct function of $\mathcal{I}(\mathbf{X})$. A design $\mathbf{X}$ is called $\mathcal{D}$-optimal if it maximizes the determinant of the information matrix, or equivalently, if it minimizes the determinant of the variance-covariance matrix of the parameter estimators. In this way, the generalized variance of the parameter estimators is minimized (Atkinson and Donev 1992). It is shown in Goos and Vandebroek (2001b) that because of the compound symmetric error structure of model (2), (16) is equal to

$$\mathcal{I}(\mathbf{X}) = \frac{1}{\sigma^2_\varepsilon}\left\{\mathbf{X}'\mathbf{X} - \sum_{i=1}^{b}\frac{\sigma^2_\gamma/\sigma^2_\varepsilon}{1+m_i(\sigma^2_\gamma/\sigma^2_\varepsilon)}(\mathbf{X}_i'\mathbf{1}_{m_i})(\mathbf{X}_i'\mathbf{1}_{m_i})'\right\}. \tag{17}$$

In terms of the coefficient of correlation (10), (17) can be rewritten as

$$\mathcal{I}(\mathbf{X}) = \frac{1}{\sigma_\varepsilon^2}\left\{\mathbf{X}'\mathbf{X} - \sum_{i=1}^{b}\frac{\rho}{1+\rho(m_i-1)}(\mathbf{X}_i'\mathbf{1}_{m_i})(\mathbf{X}_i'\mathbf{1}_{m_i})'\right\}. \tag{18}$$

If $\rho \to 0$, then the information matrix becomes

$$\mathcal{I}(\mathbf{X}) = \frac{1}{\sigma_\varepsilon^2}\mathbf{X}'\mathbf{X}, \tag{19}$$

which is the information matrix for the uncorrelated model (11). We refer to a design that maximizes the $\mathcal{D}$-criterion value $|\sigma_\varepsilon^{-2}\mathbf{X}'\mathbf{X}|$ as a $\mathcal{D}$-optimal completely randomized design (CRD). In contrast, a design that maximizes the $\mathcal{D}$-criterion value $|\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}|$ for ρ significantly differing from zero is called a $\mathcal{D}$-optimal conjoint design. It follows from the determinant expressions in these definitions that the relative statistical efficiency of a CRD is not affected by σ_ε^2 whereas that of a conjoint design depends on ρ through $\mathbf{V}$. Note that a CRD can be seen as a conjoint design in which each of the profiles is assigned to a different respondent so that as many respondents as design profiles are needed. In that case, $\mathbf{V}$ is a diagonal matrix and the variability within respondents, σ_ε^2, cannot be distinguished from the variability between respondents, σ_γ^2.

If $\rho \to 1$, then (18) results in the information matrix for the model with fixed respondent effects. When respondent effects are fixed, or non-stochastic, interest lies in the effects of the individual respondents and not in the possible effects of the population where the respondents belong to. In practice, the finding implies the $\mathcal{D}$-optimal design in the presence of random respondent effects is equivalent to the $\mathcal{D}$-optimal design in the presence of fixed respondent effects.

In our study in Section 5, we first verify whether $\mathcal{D}$-optimal conjoint designs are statistically more efficient than $\mathcal{D}$-optimal CRDs. To that end, we compare the $\mathcal{D}$-criterion values of the designs. A necessary condition to compare these values is that the variability assumed in the designs is the same. For that purpose, we assume without loss of generality a total variance of one in the designs. Hence, setting the only variance component σ_ε^2 to one in the information matrix (19), the $\mathcal{D}$-criterion value of a $\mathcal{D}$-optimal CRD becomes

$$|\mathbf{X}'\mathbf{X}|. \tag{20}$$

However, to compute the proper $\mathcal{D}$-criterion value of a $\mathcal{D}$-optimal conjoint design, we have to take into account two variance components, σ_ε^2 and σ_γ^2. These components have to sum to one so that $\sigma_\varepsilon^2 = 1-\rho$ and the $\mathcal{D}$-criterion value of a $\mathcal{D}$-optimal conjoint design comes down to

$$(1-\rho)^{-k}|\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}|, \tag{21}$$

with k the number of parameters. Usually, $\mathcal{D}$-criterion values are expressed per parameter so that the $\mathcal{D}$-criterion values of a $\mathcal{D}$-optimal CRD and a $\mathcal{D}$-optimal conjoint design amount to

$$|\mathbf{X}'\mathbf{X}|^{1/k} \text{ and } (1-\rho)^{-1}|\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}|^{1/k}. \tag{22}$$

4 Design construction algorithm

The algorithm to generate $\mathcal{D}$-optimal designs for conjoint experiments is adapted from the algorithm of Goos and Vandebroek (2004) for the construction of $\mathcal{D}$-optimal split-plot designs. Split-plot designs are similar to block designs in that they group observations together, but the underlying reason for the grouping is different. In a split-plot design, the grouping occurs on the basis of common factor levels for some of the experimental factors which are, in some sense, hard to change. In contrast, the grouping in a block design depends on a certain characteristic of the observations independent of the factors under investigation. In our conjoint setting, this characteristic refers to the fact that each block of profiles is evaluated by a different respondent.

The algorithm to construct $\mathcal{D}$-optimal conjoint designs starts with the composition of the set of candidate profiles. For example, for one type of designs in our study in Section 5, we used four attributes, each acting at three levels so that the set of candidates consists of $3^4 = 81$ profiles in that case. Next, a starting design is computed by first randomly selecting a number of profiles from the list of candidates. The first selected profile constitutes the first block after which each of the other profiles is randomly assigned to an existing block or to a new block. The starting design is completed by using a greedy heuristic that sequentially adds the candidate profile that produces the largest increase in the determinant. Also in this case, the profiles are randomly assigned to an existing or to a new block.

In order to improve the starting design, two procedures are applied consecutively, namely interchanging design profiles from different blocks and exchanging design profiles with candidate profiles. In the interchange procedure, all possible interchanges of design profiles from different blocks are evaluated, but only the best one is carried out. This process is repeated until no further improvement can be made. In the exchange procedure, three alternative exchange strategies are considered for each combination of a design profile and a candidate. In each of the strategies, the design profile is removed from the design and the candidate is added to the design. First, the candidate entering the design can be assigned to the same block as the profile removed from the design. Second, the entering profile can be assigned to another block than that of the removed profile. Finally, the entering profile can also be assigned to a new block. When all possible exchanges have been evaluated, the best one is performed. This procedure is repeated until improvement stops. The algorithm then returns to the interchange procedure and continues with evaluating interchanges and exchanges until no better design can be obtained. To avoid being stuck in a locally optimal design, more than one starting design is generated and the design search is repeated. Each repetition is called a try and the most efficient design from all tries is referred to as the $\mathcal{D}$-optimal conjoint design.

It is clear that the design problem of finding the optimal number of respondents, derived from the number of blocks the $\mathcal{D}$-optimal conjoint design contains, and selecting the profiles for each respondent is very complex. The more observations, attributes and attribute levels are involved, the more designs are possible and the higher the chance the algorithm yields poor local optima. Therefore, for large problem situations, a great

number of tries is needed. For example, we used 6000 tries to compute $\mathcal{D}$-optimal conjoint designs consisting of 60 profiles in the case of four attributes, each acting at three levels. Nevertheless, we could benchmark the results by means of the algorithm of Goos and Vandebroek (2001a) for the construction of optimal designs in the presence of random block effects. We refer to this algorithm as the benchmark algorithm and to the adapted version of the algorithm of Goos and Vandebroek (2004) as the main algorithm. The benchmark algorithm proceeds largely in the same way as the main algorithm. However, it differs in one crucial aspect from the main algorithm. Whereas the main algorithm determines the optimal number of respondents needed, b, and the optimal number of profiles for each respondent, $m_1, ..., m_b$, the benchmark algorithm requires the values b and $m_1, ..., m_b$ as an input. In other words, the benchmark algorithm computes the $\mathcal{D}$-optimal conjoint design for a given design structure, represented by b and $m_1, ..., m_b$, which is not necessarily the optimal design structure. Only the main algorithm is developed in such a way to provide the optimal design structure. As a consequence, the exchange procedure is limited to replacing the design profile by the candidate in the same block. Moreover, it is obvious that the design space the benchmark algorithm has to explore is much more constrained than that of the main algorithm. As a result, the benchmark algorithm requires less tries and computation time.

Finally, to use the main and benchmark algorithms, an estimate of ρ must be provided. As a first illustration, we calculated an estimate $\hat{\rho}$ using the dataset from a ratings-based conjoint experiment carried out in the health economics area and described in Brazier et al. (2002). The dataset contains the ratings of 611 respondents who each had to evaluate a different set of six health states. By means of the SAS procedure PROC MIXED and REML as its default estimation method to estimate model (2), we obtained a value of $\hat{\rho} = 0.62$. As a second example, we received a dataset from a sensory experiment performed by the multinational brewer InBev. The experiment consisted of 100 respondents who each had to rate four types of beer with respect to four sensory characteristics. The estimates $\hat{\rho}$ for each of these characteristics amount to 0.48, 0.46, 0.36 and 0.41, respectively. In Section 5.3, we show however that for proper design construction the value of ρ does not need to be known as $\mathcal{D}$-optimal conjoint designs are not very sensitive to misspecification of ρ.

5 Results

In this section, we present a selection of computational results that are representative of $\mathcal{D}$-optimal conjoint designs of any type. We first prove that it is statistically justified to apply these designs instead of $\mathcal{D}$-optimal CRDs. We then proceed with a discussion of the optimal design structure and the computation times needed. Next, we deal with some practical issues and seek for ways to save on computation time. Lastly, $\mathcal{D}$-optimal arrangements of all candidate profiles are studied.

5.1 Design selection

We computed $\mathcal{D}$-optimal conjoint designs for 11 different sample sizes: 10, 20, 24, 30, 36, 40, 50, 60, 70, 72 and 81. The profiles are composed of four attributes, each acting at a certain level. We investigated two possible settings of the attribute levels. In the first setting, all attributes have three levels and thus the attribute levels are homogeneous. In contrast, in the second setting, the attribute levels are heterogeneous: the first attribute acts at two levels, the second attribute at three levels, the third attribute at four levels and the fourth attribute again at three levels. As a consequence, the number of candidate profiles amounts to $3^4 = 81$ in the homogeneous level setting and to $2 \times 3 \times 4 \times 3 = 72$ in the heterogeneous level setting. For each of the design problems, we applied 10 values of the degree of correlation $\rho : 0, 0.1, ..., 0.9$. We included $\rho = 0$ to compute the $\mathcal{D}$-optimal CRDs. Hence, in total, we constructed $11 \times 2 \times 9 = 198$ $\mathcal{D}$-optimal conjoint designs and $11 \times 2 = 22$ $\mathcal{D}$-optimal CRDs.

5.2 Comparison of $\mathcal{D}$-criterion values

In Table 1, the $\mathcal{D}$-criterion values of the $\mathcal{D}$-optimal conjoint designs and CRDs are displayed. As discussed in Section 3, these values were calculated by means of the expressions in (22) to compare the $\mathcal{D}$-optimal conjoint designs with the $\mathcal{D}$-optimal CRDs for the same sample size and setting of the attribute levels. It turns out that the $\mathcal{D}$-optimal CRDs are outperformed by each of the corresponding $\mathcal{D}$-optimal conjoint designs. This result was expected as Goos (2002), page 133, extensively proved that $\mathcal{D}$-optimal block designs are more efficient than $\mathcal{D}$-optimal CRDs provided the experimental situation exhibits a block format. (see also Section 1). Hence, since profile ratings from the same respondent are correlated ($\rho \neq 0$), it is statistically justified to take into account the compound symmetric error structure when designing conjoint experiments. In addition, Table 1 reveals that the higher the correlation, the larger the efficiency gain of using a $\mathcal{D}$-optimal conjoint design instead of a $\mathcal{D}$-optimal CRD. Also this result is not a surprise as Goos (2002), page 133, noted a similar finding for $\mathcal{D}$-optimal block designs.

It is obvious that the $\mathcal{D}$-criterion values increase with the sample size because more information on the parameters is acquired by enlarging the experiment. It is however conceivable that a saturation level of $\mathcal{D}$-efficiency is reached when exceeding a certain number of observations. For that reason, we examined more thoroughly the relationship between the sample size and the $\mathcal{D}$-criterion values by plotting them against each other for the 10 degrees of correlation. The plot for the homogeneous level setting is displayed in Figure 1. The plot for the heterogeneous level setting exhibits the same pattern and is therefore not depicted. The plots show that the $\mathcal{D}$-criterion values increase linearly with the sample size, hereby contracting the idea of a saturation level of $\mathcal{D}$-efficiency for the sample sizes considered.

5.3 Optimal number of respondents and profile selection

In the previous section, we demonstrated that the $\mathcal{D}$-optimal conjoint designs are statistically much more efficient than the $\mathcal{D}$-optimal CRDs. As a result, the random respondent

Table 1: $\mathcal{D}$-criterion values of the $\mathcal{D}$-optimal conjoint designs for 11 different sample sizes n, for homogeneous and heterogeneous attribute levels and for various degrees of correlation ρ.

	hom/	ρ									
n	het	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
10	hom	5.965	6.205	6.640	7.271	8.153	9.409	11.290	14.376	20.356	37.271
10	het	5.311	5.575	5.996	6.586	7.402	8.556	10.279	13.102	18.569	34.022
20	hom	12.088	12.957	14.112	15.628	17.661	20.497	24.699	31.553	44.799	82.200
20	het	11.247	11.952	12.937	14.263	16.068	18.603	22.376	28.543	40.478	74.200
24	hom	14.506	15.537	16.959	18.815	21.297	24.753	29.865	38.196	54.285	99.699
24	het	13.768	14.632	15.826	17.435	19.627	22.710	27.299	34.807	49.338	90.407
30	hom	18.253	19.578	21.369	23.708	26.836	31.190	37.632	48.130	68.404	125.628
30	het	17.088	18.158	19.637	21.632	24.349	28.171	33.862	43.189	61.249	112.283
36	hom	22.093	23.775	25.951	28.791	32.590	37.878	45.701	58.450	83.071	152.564
36	het	20.696	22.007	23.810	26.237	29.538	34.181	41.090	52.392	74.266	136.086
40	hom	24.428	26.187	28.535	31.611	35.755	41.532	50.087	64.036	90.982	167.051
40	het	22.908	24.315	26.297	28.973	32.617	37.762	45.414	57.927	82.141	150.566
50	hom	30.583	32.833	35.810	39.701	44.913	52.174	62.921	80.442	114.286	209.826
50	het	28.709	30.500	32.986	36.345	40.919	47.353	56.932	72.598	102.920	188.612
60	hom	36.737	39.497	43.112	47.830	54.141	62.925	75.922	97.101	138.002	253.450
60	het	34.505	36.682	39.679	43.716	49.212	56.940	68.448	87.273	123.709	226.684
70	hom	42.895	46.108	50.277	55.728	63.052	73.256	88.363	112.988	160.551	294.818
70	het	40.238	42.758	46.243	50.947	57.354	66.365	79.781	101.726	144.200	264.238
72	hom	44.185	47.514	51.862	57.538	65.129	75.697	91.331	116.809	166.013	304.893
72	het	41.449	44.051	47.643	52.489	59.087	68.368	82.186	104.787	148.534	272.173
81	hom	49.709	53.494	58.390	64.780	73.327	85.225	102.827	131.512	186.909	343.270
81	het	46.577	49.503	53.542	58.987	66.403	76.832	92.361	117.761	166.925	305.873

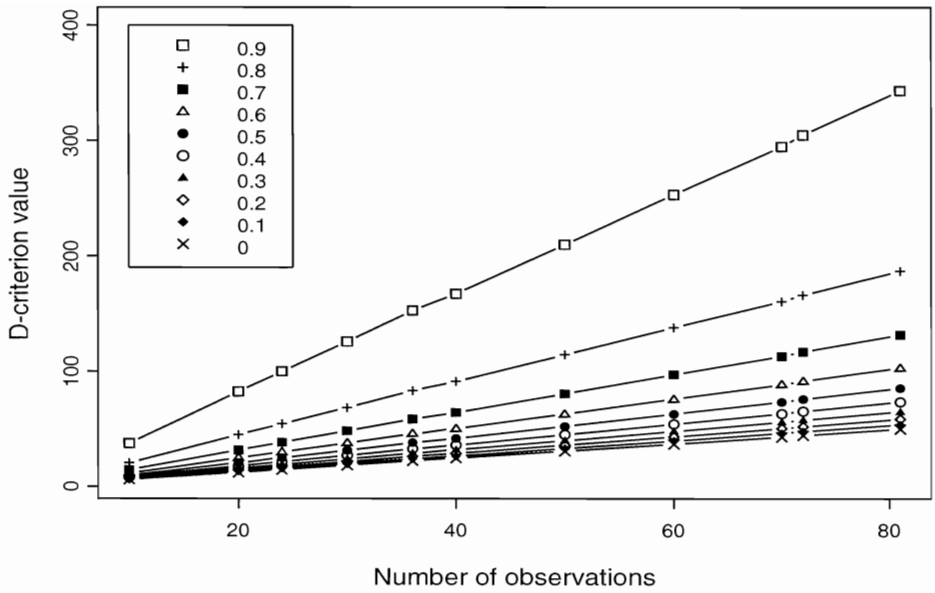

Figure 1: $\mathcal{D}$-criterion values of the $\mathcal{D}$-optimal conjoint designs for sample sizes ranging from 10 to 81, for homogeneous attribute levels and for degrees of correlation of $0, 0.1, ..., 0.9$.

effects model (2) is appropriate and it is prudent to assign a block of profiles to each of the respondents. Note that if the reverse was true, the uncorrelated model (11) had to be used and each respondent had to evaluate only one profile. In this section, we focus therefore on the $\mathcal{D}$-optimal conjoint designs to derive more precisely how many respondents are required for a specific conjoint setting and which and how many profiles are offered to each of them. From the designs, we easily observe how many respondents are needed for a specific setting through the number of blocks the designs contain. The profiles in a block are then assigned to one respondent. In Table 2, the blocking or respondent-profile structures of the 198 $\mathcal{D}$-optimal conjoint designs are presented. The designs themselves are not portrayed, but can be obtained from the authors.

In total, we found 26 different $\mathcal{D}$-optimal conjoint designs. This means that most of the designs for a particular sample size and setting of the attribute levels are optimal for a broad range of degrees of correlation. Consequently, the $\mathcal{D}$-optimal conjoint designs are fairly robust against misspecifications of the degree of correlation. Also Goos (2002), page 122, observed this result while computing $\mathcal{D}$-optimal block designs for several degrees of correlation.

It turns out it is often statistically most efficient to administer three profiles to respondents. Nevertheless, when a heterogeneous number of levels is used and/or the degree of correlation is increased, it is sometimes most efficient to administer four profiles to one or more respondents. In that case, the optimal number of respondents decreases. The result

Table 2: Respondent-profile structures of the $\mathcal{D}$-optimal conjoint designs for 11 different sample sizes n, for homogeneous and heterogeneous attribute levels and for 9 degrees of correlation $\rho \in \{0.1; 0.2; ...; 0.9\}$.

Design	n	hom/het	ρ	Respondent-profile structure	b
1	10	hom	$\{0.1; ...; 0.9\}$	$m_1 = 3 \mid m_2 = 7$	2
2	10	het		$m_1 = 4 \mid m_2 = 6$	2
3	20	hom	$\{0.1; ...; 0.9\}$	$m_1, ..., m_6 = 3 \mid m_7 = 2$	7
4	20	het		$m_1, ..., m_4 = 3 \mid m_5, m_6 = 4$	6
5	24	hom	$\{0.1; ...; 0.9\}$	$m_1, ..., m_8 = 3$	8
6	24	het			
7	30	hom	$\{0.1; ...; 0.9\}$	$m_1, ..., m_{10} = 3$	10
8	30	het	$\{0.1; ...; 0.6\}$	$m_1, ..., m_{10} = 3$	10
9	30		$\{0.7; 0.8; 0.9\}$	$m_1, ..., m_6 = 3 \mid m_7, m_8, m_9 = 4$	9
10	36	hom	$\{0.1; ...; 0.9\}$	$m_1, ..., m_{12} = 3$	12
11	36	het			
12	40	hom	$\{0.1; 0.2\}$	$m_1, ..., m_{13} = 3 \mid m_{14} = 1$	14
13	40		$\{0.3; ...; 0.9\}$	$m_1, ..., m_{12} = 3 \mid m_{13} = 4$	13
14	40	het	$\{0.1; ...; 0.4\}$	$m_1, ..., m_{12} = 3 \mid m_{13} = 4$	13
15	40		$\{0.5; ...; 0.9\}$	$m_1, ..., m_8 = 3 \mid m_9, ..., m_{12} = 4$	12
16	50	hom	$\{0.1; ...; 0.9\}$	$m_1, ..., m_{16} = 3 \mid m_{17} = 2$	17
17	50	het		$m_1, ..., m_{14} = 3 \mid m_{15}, m_{16} = 4$	16
18	60	hom	$\{0.1; ...; 0.9\}$	$m_1, ..., m_{20} = 3$	20
19	60	het			
20	70	hom	$\{0.1; 0.2\}$	$m_1, ..., m_{23} = 3 \mid m_{24} = 1$	24
21	70		$\{0.3; ...; 0.9\}$	$m_1, ..., m_{22} = 3 \mid m_{23} = 4$	23
22	70	het	$\{0.1; ...; 0.9\}$	$m_1, ..., m_{22} = 3 \mid m_{23} = 4$	23
23	72	hom	$\{0.1; ...; 0.9\}$	$m_1, ..., m_{24} = 3$	24
24	72	het			
25	81	hom	$\{0.1; ...; 0.9\}$	$m_1, ..., m_{27} = 3$	27
26	81	het			

that sometimes more profiles are grouped when the correlation is increased is related to Goos and Vandebroek (2004) who found that it generally holds for split-plot designs to group more experimental runs when the correlation is higher.

Finally, the case of 10 observations is somewhat special in the sense that the profiles are not so equally spread over respondents in blocks of three and four as in the other design cases. Instead, two respondents are required to evaluate three and seven profiles, respectively, in the homogeneous level setting and four and six profiles in the heterogeneous level setting. As a consequence, assigning five profiles to two respondents each is also not an optimal design option. This is a striking result which is similar to Goos (2002), page 133, who observed that the two blocks of a 10 run $\mathcal{D}$-optimal block design for quantitative factors and a second order model are asymmetric.

5.4 Computation time and number of tries

This section presents an overview of the computation times per 1000 tries and the numbers of tries utilized to generate the $\mathcal{D}$-optimal conjoint designs by means of the main algorithm as well as the benchmark algorithm. Recall from Section 4 that the main algorithm is vital for determining the optimal respondent-profile structures whereas the benchmark algorithm requires a respondent-profile structure as an input to generate the $\mathcal{D}$-optimal conjoint design corresponding to that structure. To register the computation times, we applied the optimal respondent-profiles structures to the benchmark algorithm, although the use of other structures, except for the CRD structure of one profile per respondent, would not affect the times by much. Both algorithms were implemented in Fortran 77. The computation times were obtained by means of a Dell PC with a 1.80 GHz Intel Processor and 256 MB RAM. In Table 3, exact times per 1000 tries are reported to construct the $\mathcal{D}$-optimal conjoint designs for $\rho = 0.5$. The times in Table 3 are however representative of the $\mathcal{D}$-optimal conjoint designs for any of the 9 values of ρ ($\rho \in \{0.1; 0.2; ...; 0.9\}$). Table 3 further lists the numbers of tries used to generate the $\mathcal{D}$-optimal conjoint designs. We believe these numbers are sufficient in order not to be stuck in locally optimal designs.

To study the computation times per 1000 tries more rigorously, we plotted them against the sample size. The plot for the homogeneous level setting is displayed in Figure 2. The plot for the heterogeneous level setting is omitted as it shows a similar pattern. Exploring the times per 1000 tries and the numbers of tries used, we observed that it takes very long to compute the $\mathcal{D}$-optimal conjoint designs with a sample size of 40 and more by means of the main algorithm. More specifically, computation times with the main algorithm grow exponentially with the sample size. On the other hand, the benchmark algorithm goes much faster and computation times with this algorithm increase slowly with the sample size. Lastly, note that for the construction of $\mathcal{D}$-optimal CRDs, we used the benchmark algorithm that, in this case, only takes a quarter of the times registered for this algorithm in Table 3. However, to generate $\mathcal{D}$-optimal CRDs, there are much faster algorithms in the literature on optimal design (e.g. BLKL-algorithm, see Atkinson and Donev 1992) which are specially developed for this purpose.

Table 3: Approximate computation times per 1000 tries and number of tries used to generate the $\mathcal{D}$-optimal conjoint designs by means of the main algorithm and the benchmark algorithm.

		Main algorithm		Benchmark algorithm	
n	hom/het	time/1000 tries	# tries	time/1000 tries	# tries
10	hom	00.21h	1000	15s	1000
10	het	00.23h	1000	15s	1000
20	hom	01.17h	2000	00.02h	1000
20	het	01.05h	2000	00.01h	1000
24	hom	02.04h	3000	00.02h	2000
24	het	01.24h	3000	00.02h	2000
30	hom	02.54h	3000	00.03h	2000
30	het	02.02h	3000	00.03h	2000
36	hom	03.58h	4000	00.04h	2000
36	het	02.47h	4000	00.03h	2000
40	hom	04.52h	4000	00.05h	2000
40	het	03.05h	4000	00.04h	2000
50	hom	08.31h	5000	00.08h	3000
50	het	04.36h	5000	00.07h	3000
60	hom	11.05h	6000	00.10h	3000
60	het	07.14h	6000	00.09h	3000
70	hom	15.42h	7000	00.15h	4000
70	het	09.18h	7000	00.15h	4000
72	hom	16.44h	7000	00.15h	4000
72	het	10.43h	7000	00.15h	4000
81	hom	21.19h	8000	00.21h	5000
81	het	12.51h	8000	00.18h	5000

5.5 Compromising between practical and optimal respondent-profile structures

In this section, we examine whether some of the optimal respondent-profile structures in Table 2 can be slightly adapted to result in more practical structures for which the corresponding $\mathcal{D}$-optimal conjoint designs are still statistically fairly efficient. These $\mathcal{D}$-optimal conjoint designs are computed by means of the benchmark algorithm given a more practical respondent-profile structure as an input. We tackle the following three cases. The first case concerns the designs in which one or two profiles are administered to one of the respondents. It may be more sensible, however, to assign four profiles instead of three to one or two respondents so that one respondent less is needed. The second case covers the designs in which the profiles are not spread out as equally as possible over respondents in blocks of three. It would be very convenient for experimenters if the respondent-profile structure of three profiles per respondent could be generalized to the instances concerned.

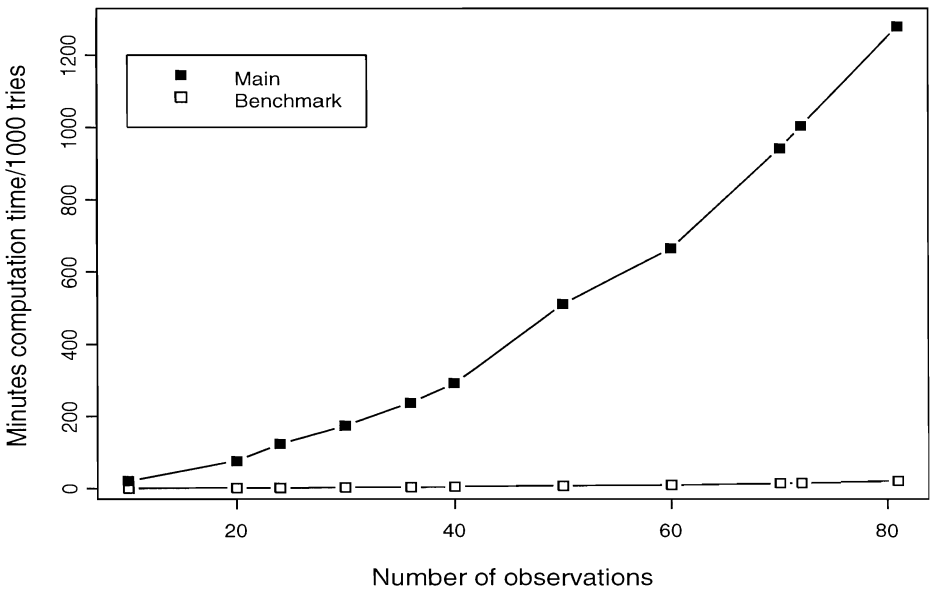

Figure 2: Minutes of computation time per 1000 tries to generate the $\mathcal{D}$-optimal conjoint designs by means of the main algorithm and the benchmark algorithm.

The third case involves in fact all the designs, except for those with 10 observations, as it may be more useful to administer more than three profiles per respondent so as to reduce the number of respondents needed. We discuss these three cases more in detail and investigate how much one looses in $\mathcal{D}$-efficiency by applying more practical respondent-profile structures than the optimal ones in Table 2. If the losses in $\mathcal{D}$-efficiency incurred are negligible, then we retain the more practical structures.

The approach we have adopted to express losses in $\mathcal{D}$-efficiency when respondent-profile structures are used different from the optimal ones in Table 2, is a functional one and proceeds as follows. For each setting of the attribute levels and degree of correlation ($\rho \neq 0$), we performed a regression analysis of the $\mathcal{D}$-optimal criterion values of Table 1 with respect to the sample size. We denote these $\mathcal{D}$-optimal criterion values as $\mathcal{D}_{n^*}^{opt}$. In total, the regression analyses yield $2 \times 9 = 18$ values for the intercept and the slope. The intercept and slope from a regression are referred to as ψ and ω. We then computed $\mathcal{D}$-optimal conjoint designs for other respondent-profile structures than the optimal ones by means of the benchmark algorithm. These designs have one of the 11 different sample sizes n and their $\mathcal{D}$-criterion values are denoted by $\mathcal{D}_n^{sub}$ since sub-optimal respondent-profile structures are used. It is clear that the $\mathcal{D}$-criterion values $\mathcal{D}_n^{sub}$ are lower than the $\mathcal{D}$-criterion values $\mathcal{D}_{n^*}^{opt}$ for $n = n^*$. Hence, for each of the $\mathcal{D}$-criterion values $\mathcal{D}_n^{sub}$, we derived how many observations n^* are required in the optimal case using the equations

$$\mathcal{D}_n^{sub} = \mathcal{D}_{n^*}^{opt} \text{ and} \tag{23}$$

$$\mathcal{D}_{n^*}^{opt} = \psi + \omega n^*. \tag{24}$$

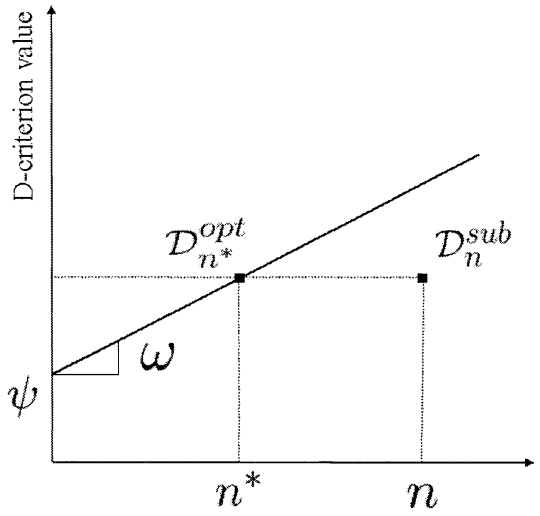

Figure 3: Derivation of the redundant numbers of observations, $n - n^*$, using regression analysis.

To illustrate, expressions (23) and (24) are set out in Figure 3. Obviously, all the sample sizes n^* found are lower than or, due to rounding, equal to the sample sizes n for which we constructed $\mathcal{D}$-optimal conjoint designs with sub-optimal respondent-profile structures. Finally, the differences in number of observations, $n - n^*$, give a clear indication of the extent of the efficiency losses when using sub-optimal respondent-profile structures. The larger the differences, the higher the losses. We refer to these differences as redundant numbers of observations as they specify how many observations would be redundant if optimal designs are applied to reach a given level of $\mathcal{D}$-efficiency.

One or two profiles for one respondent

For sample sizes of 40 and 70 in the homogeneous level setting for values of ρ of 0.1 and 0.2, it is statistically most efficient to administer one profile to one of the respondents. The assignment of one profile to one respondent seems however not very efficient from a practical point of view. Moreover, if there are costs associated per respondent needed, then it seems also not very cost efficient. If the profiles are easy to rate, then the same reasoning holds for the designs in which two profiles are administered to one of the respondents. These are the designs with a sample size of 20 and 50 in the homogeneous level setting. For these two design cases, we calculated how much one looses in $\mathcal{D}$-efficiency when assigning four profiles instead of three to one or two respondents in order to save on one respondent. It turns out that the losses in $\mathcal{D}$-efficiency from applying the more practical respondent-profile structures are so small that there are no redundant numbers of observations. As a result, the structures constitute a good compromise between practical and statistical efficiency.

Blocks of three profiles

As discussed in Section 5.3, blocks of three profiles are most often assigned to respondents. Nevertheless, when a heterogeneous number of levels is used and/or the degree of correlation is increased, blocks of four profiles are sometimes administered. In the case of 10 observations, even a block of seven profiles is administered in the homogeneous level setting and one of six profiles in the heterogeneous level setting. For these particular cases, we examined whether we can more equally spread out the profiles over respondents in blocks of three. For sample sizes that are not multiples of three, we opted for one or two blocks of four profiles (see the previous section). The reason for this study is that if the efficiency losses incurred turn out to be negligible, we can propose the general respondent-profile structure of three profiles per respondent. Such a general structure allows to apply the benchmark algorithm to compute $\mathcal{D}$-optimal conjoint designs for all possible sample sizes. As shown in Section 5.4, the benchmark algorithm drastically speeds up the computation time. On the other hand, one may have to designate more respondents than optimally needed.

Apart from the optimal designs with 10 observations, we did not find any redundant numbers of observations when more equally dividing the profiles over respondents in blocks of three in the cases concerned. For a sample size of 10, the redundant numbers of observations are listed in Table 4. In part a, redundant numbers of observations are given for the respondent-profile structure that assigns three profiles to three respondents each and one profile to a fourth respondent. In part b on the other hand, they are shown for the more practical structure of three profiles to two respondents each and four profiles to a third respondent. In general, the redundant numbers are relatively high since the difference between the optimal respondent-profile structures and the structures in Table 4 is quite large. The redundant numbers of observations are sometimes less for the more practical structure in part b as this structure resembles the optimal one more. The redundant numbers of observations further increase with the degree of correlation. Significant differences between the homogeneous and heterogeneous level settings are however not observed.

Table 4: Redundant numbers of observations indicating the degree of sub-optimality in case 10 profiles are divided in a statistically efficient way over a) four respondents in three blocks of three and one block of one; b) three respondents in two blocks of three and one block of four.

		hom/		ρ								
	n	het	b	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
a)	10	hom	4	0	1	1	1	2	2	3	3	4
	10	het	4	1	1	1	2	2	2	3	3	4
b)	10	hom	3	0	0	1	1	1	2	2	2	3
	10	het	3	1	1	1	1	1	2	2	2	3

It follows from the account in this section that we can generalize the concept of blocks of three profiles to all design cases considered except for the case of 10 observations. To investigate from which sample size we may start applying blocks of three profiles, we computed $\mathcal{D}$-optimal conjoint designs that equally spread the profiles over respondents in blocks of three for sample sizes between 11 and 20. We again chose for one or two blocks of four profiles in case the sample size is not a multiple of three. It turns out that the designs with a sample size of 11 and more do not result in any redundant numbers of observations. Hence, for sample sizes larger than or equal to 11, we propose the rule of thumb of three profiles per respondent. A consequence of this rule is that in case there are respondent costs, instead of first deciding on the number of observations, it may be more sensible to first decide on the number of respondents to appoint. If each respondent is then given three profiles to rate, the sample size is easily found.

Blocks of more than three profiles

A disadvantage of applying the general rule of three profiles per respondent is that it requires relatively many respondents. Therefore, when profiles are very easy to rate, it may be more practical to administer more than three profiles per respondent. To that end, we investigated several combinations of numbers of respondents b and numbers of profiles m assigned to each of the respondents given a sample size n. For example, if $n = 72$, then it is interesting to examine the combinations $b = 18$, $m = 4$; $b = 12$, $m = 6$; $b = 9$, $m = 8$ and $b = 8$, $m = 9$. For each of these combinations, we calculated the redundant numbers of observations to find the combination with the smallest losses in $\mathcal{D}$-efficiency. We considered the 10 sample sizes starting at 20, the homogeneous and heterogeneous level settings and the 9 degrees of correlation. We expected the redundant numbers of observations to decrease with the number of respondents b or to increase with the number of profiles m per respondent.

It turns out that our expectation indeed holds for the $\mathcal{D}$-optimal conjoint designs in the heterogeneous level setting that correspond to several combinations of b and m given n. However, in the homogeneous level setting, the result is surprisingly different. In this setting, the $\mathcal{D}$-optimal conjoint designs in which the number of profiles per respondent is a multiple of three, represent the smallest efficiency losses. This is illustrated in Table 5 where we have listed the redundant numbers of observations for sample sizes of 60 and 72. For 60 observations, the redundant numbers of observations from administering six profiles to ten respondents ($b = 10, m = 6$) are lower than or equal to those from administering five profiles to twelve respondents ($b = 12, m = 5$) or four profiles to fifteen respondents ($b = 15, m = 4$). For 72 observations, we observe that the redundant numbers of observations from assigning nine profiles to eight respondents ($b = 8, m = 9$) and six profiles to twelve respondents ($b = 12, m = 6$) are lower than or equal to those from assigning eight profiles to nine respondents ($b = 9, m = 8$) and four profiles to eighteen respondents ($b = 18, m = 4$), respectively. The finding that the efficiency losses are smallest when multiples of three profiles are administered in the homogeneous level setting is also valid for sample sizes that are not nicely divisible by three. For example, in the case of 40 observations, the redundant numbers of observations from assigning six profiles to six

respondents and four profiles to a seventh respondent are smaller than or equal to those from assigning four profiles to ten respondents. To conclude, for all the design cases considered, we found that the redundant numbers of observations increase with the degree of correlation.

Table 5: Redundant numbers of observations indicating the degree of sub-optimality in case 60 and 72 profiles in the homogeneous level setting are equally spread out in a statistically efficient way over different numbers of respondents b.

			ρ								
n	b	m	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
60	6	10	3	5	6	6	7	7	7	7	8
	10	6	1	2	3	3	4	4	4	4	4
	12	5	2	3	3	4	4	4	5	5	5
	15	4	1	2	3	4	4	4	4	5	5
72	8	9	3	5	6	6	7	7	7	8	8
	9	8	3	5	6	6	7	7	8	8	8
	12	6	2	3	3	4	4	4	5	5	5
	18	4	2	3	4	4	5	5	5	6	6

5.6 Replicating $\mathcal{D}$-optimal conjoint designs

In this section, we investigate whether replicating small $\mathcal{D}$-optimal conjoint designs is statistically as efficient as generating larger $\mathcal{D}$-optimal conjoint designs. Clearly, less computation time would be needed if a large conjoint experiment could be set up by means of a small design. To evaluate the statistical efficiency of design plans that contain replications of small $\mathcal{D}$-optimal conjoint designs, we calculated the redundant numbers of observations associated with these plans. These numbers are derived by first computing the $\mathcal{D}$-criterion values of the plans. The $\mathcal{D}$-criterion value, $\mathcal{D}_n$, of a design plan consisting of a multiple c of a small $\mathcal{D}$-optimal conjoint design with sample size n_s is given by

$$\mathcal{D}_n = c \times \mathcal{D}^*_{n_s}, \tag{25}$$

where $\mathcal{D}^*_{n_s}$ denotes the $\mathcal{D}$-criterion value of the small $\mathcal{D}$-optimal conjoint design. This value is found in Table 1. Using $\mathcal{D}_n = \mathcal{D}^{opt}_{n^*}$ in combination with regression expression (24) for each design plan then yields the redundant numbers of observations, $n - n^*$.

In Table 6, redundant numbers of observations are given in the case $\mathcal{D}$-optimal conjoint designs with sample sizes of 10, 20, 24, 30 and 36 are replicated to carry out experiments with 20, 30, 40, 50, 60 and 72 observations. The numbers are only reported for the homogeneous level setting as the numbers for the heterogeneous level setting are comparable. It turns out that there are no or little redundant numbers of observations when the $\mathcal{D}$-optimal conjoint designs with sample sizes of 20, 24, 30 and 36 are replicated. On

the other hand, replicating the $\mathcal{D}$-optimal conjoint design with 10 observations results in redundant numbers of observations that increase with the degree of correlation and the number of replications.

The findings in Table 6 can be explained by two reasons. The first reason is that the optimal respondent-profile structures in the case of 10 observations differ substantially from the optimal structures in the case of 20, 24, 30 and 36 observations. The second reason comes down to the fact that maximum 10 different profiles are included for a sample size of 10 whereas more different profiles are taken up for the other sample sizes. As a result, to set up large conjoint experiments in a time-efficient manner, it is prudent to replicate the $\mathcal{D}$-optimal conjoint designs with sample sizes of 20, 24, 30 and 36, but not those with a sample size of 10.

Table 6: Redundant numbers of observations indicating the degree of sub-optimality when replications of the $\mathcal{D}$-optimal conjoint designs with sample sizes of 10, 20, 24, 30 and/or 36 are used instead of the $\mathcal{D}$-optimal conjoint designs with sample sizes of 20, 30, 40, 50, 60, 70 and 72. Numbers are given for the homogeneous level setting.

		ρ								
n	$c \times n_s$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
20	2×10	1	1	1	1	2	2	2	2	2
30	3×10	1	2	2	2	3	3	3	3	3
40	4×10	2	3	3	4	4	4	4	4	4
	2×20	0	1	1	1	1	1	1	1	1
50	5×10	3	4	4	5	5	5	5	5	6
60	6×10	3	5	5	6	6	6	7	7	7
	3×20	1	1	1	1	1	1	1	2	2
	2×30	1	1	1	0	0	0	0	0	0
70	7×10	4	5	6	7	7	8	8	8	8
72	3×24	1	1	1	1	1	1	1	1	1
	2×36	0	0	0	0	0	0	0	0	0

5.7 Randomly distributing profiles from $\mathcal{D}$-optimal CRDs

In this section, we examine how well the 22 $\mathcal{D}$-optimal CRDs perform when the profiles are randomly distributed over respondents according to the optimal respondent-profile structures in Table 2. In other words, we analyze each of the $\mathcal{D}$-optimal CRDs by means of the random respondent effects model (2) using the optimal number of respondents b and the optimal number of profiles per respondent, $m_1, ..., m_b$, for the design problem considered. If the $\mathcal{D}$-optimal CRDs perform well in this way, then they may be preferred to $\mathcal{D}$-optimal conjoint designs to save on computation time. The method we have adopted to score the $\mathcal{D}$-optimal CRDs on the random respondent effects model proceeds as follows. For each of the $\mathcal{D}$-optimal CRDs and corresponding optimal respondent-profile structure(s), we first

randomly generated 1000 arrangements of the CRD-profiles. As there are 21 different respondent-profile structures in Table 2, we obtained in total 21 arrays of 1000 profile arrangements. We then computed the $\mathcal{D}$-criterion values of all profile arrangements using the degrees of correlation associated with the respondent-profile structures.

For each array of 1000 profile arrangements, we found that the $\mathcal{D}$-criterion values for each of the corresponding degrees of correlation are bell-shaped. The average $\mathcal{D}$-criterion values for the 9 degrees of correlation in the homogeneous level setting are shown in Figure 4. For the heterogeneous level setting, a similar figure is obtained and is therefore omitted. It is interesting to compare this figure with Figure 1 in which the $\mathcal{D}$-criterion values of the $\mathcal{D}$-optimal conjoint designs are shown. As in Figure 1, the average $\mathcal{D}$-criterion values of the profile arrangements also increase with the sample size and the degree of correlation. The average $\mathcal{D}$-criterion values are however much lower than the $\mathcal{D}$-criterion values of the $\mathcal{D}$-optimal conjoint designs, suggesting that randomly distributing CRD-profiles over respondents is statistically not very efficient. In addition to the averages, we also computed the standard deviations of the $\mathcal{D}$-criterion values for each array of 1000 profile arrangements and corresponding degrees of correlation. It turns out that also the standard deviations increase with the sample size and the degree of correlation.

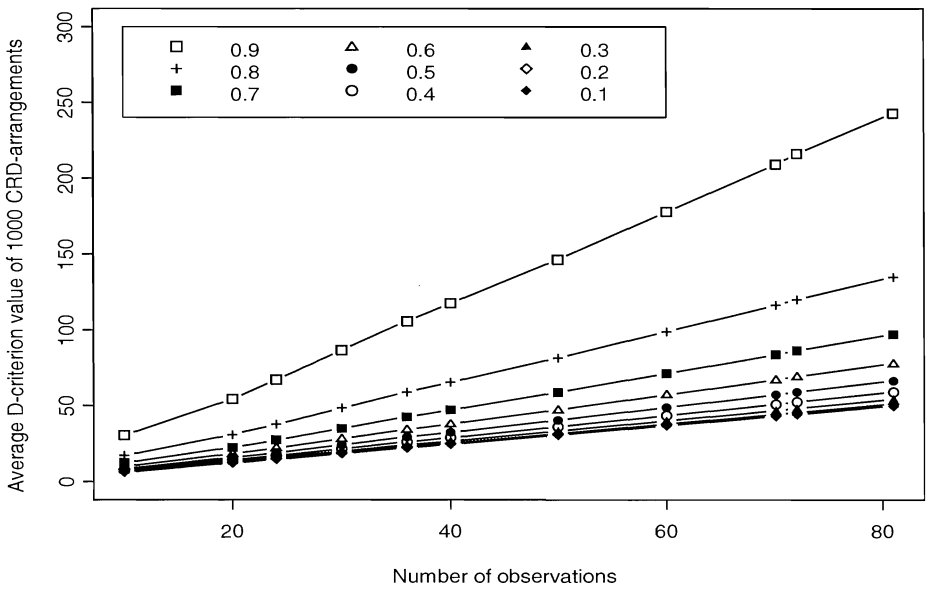

Figure 4: Average $\mathcal{D}$-criterion values of 1000 profile arrangements of the $\mathcal{D}$-optimal CRDs in the homogeneous level setting. Profiles are arranged according to the optimal respondent-profile structures.

For the average $\mathcal{D}$-criterion values, which we refer to as $\mathcal{D}_n^{avg}$, we computed the redundant numbers of observations to specify the efficiency losses from randomly distributing CRD-profiles over respondents. The numbers are derived using $\mathcal{D}_n^{avg} = \mathcal{D}_{n^*}^{opt}$ in combi-

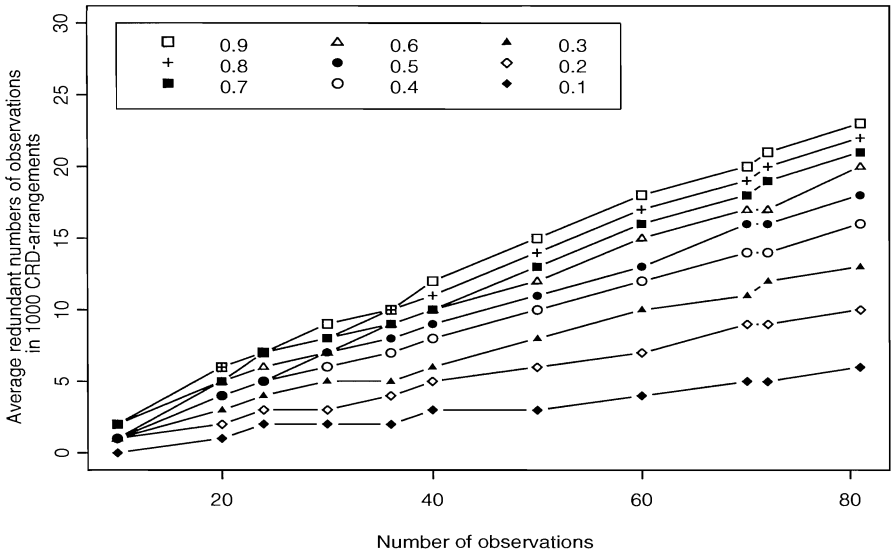

Figure 5: Average redundant numbers of observations of 1000 profile arrangements of the $\mathcal{D}$-optimal CRDs in the homogeneous level setting. Profiles are arranged according to the optimal respondent-profile structures.

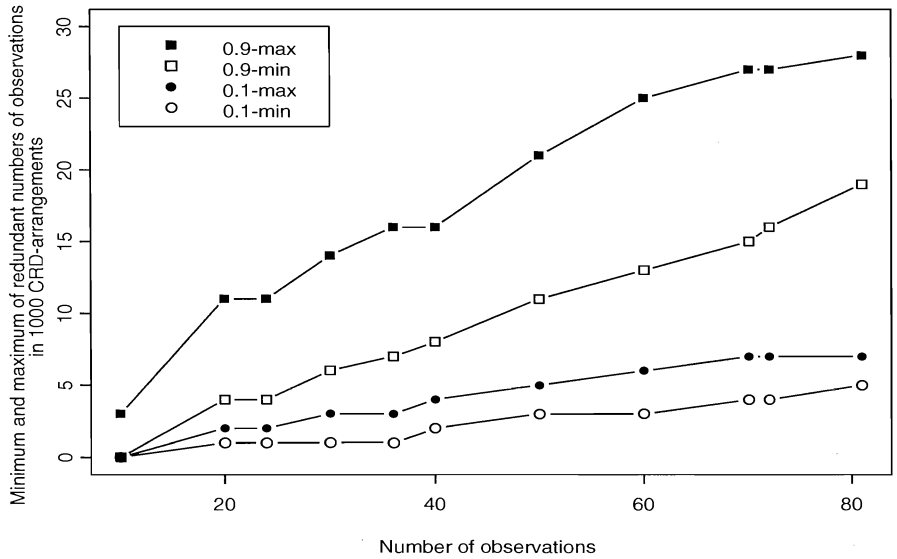

Figure 6: Minima and maxima of redundant numbers of observations of 1000 profile arrangements of the $\mathcal{D}$-optimal CRDs in the homogeneous level setting. Profiles are arranged according to the optimal respondent-profile structures for degrees of correlation of 0.1 and 0.9.

nation with regression expression (24). Obviously, these redundant numbers are average numbers. The average redundant numbers of observations in the homogeneous level setting are shown in Figure 5. The averages in the heterogeneous level setting are not displayed as they exhibit the same pattern. As expected from Figure 4, the average redundant numbers of observations are substantial and increase with the sample size and the degree of correlation.

Besides the average $\mathcal{D}$-criterion values, we also calculated the minimum and maximum $\mathcal{D}$-criterion values and corresponding redundant numbers of observations. The minimum $\mathcal{D}$-criterion values result in the largest redundant numbers of observations whereas the maximum $\mathcal{D}$-criterion values result in the smallest numbers. The maxima and minima of redundant numbers of observations in the homogeneous level setting for degrees of correlation of 0.1 and 0.9 are displayed in Figure 6. They serve as bounds between which the maxima and minima of the numbers for the other 7 degrees of correlation are comprised. For the heterogeneous level setting, the figure is practically the same and is left out. We observe that the maxima are more dispersed than the minima and that the spread between the maxima and minima increases with the degree of correlation.

5.8 Efficiently distributing the candidate profiles

This last section looks more closely at the 81 candidate profiles in the homogeneous level setting and the 72 candidate profiles in the heterogeneous level setting. For each set of candidates, we arranged the candidate profiles according to the optimal respondent-profile structure of three profiles per respondent. Numerous arrangements are possible, but we retained the arrangement with the highest $\mathcal{D}$-criterion value for both sets. Both arrangements are statistically most efficient for any of the 9 degrees of correlation. We then investigated whether the statistical efficiency of the arrangements is the same as that of the $\mathcal{D}$-optimal conjoint designs in Table 1 for the corresponding sample size and level setting. Note that the $\mathcal{D}$-optimal conjoint designs in Table 1 contain some of the candidate profiles more than once and thus some other candidates are not included. Recall further that the $\mathcal{D}$-optimal conjoint designs with sample sizes of 81 and 72 in the homogeneous and heterogenous level setting, respectively, are equivalent for the 9 degrees of correlation.

Comparing the $\mathcal{D}$-criterion values of the $\mathcal{D}$-optimal arrangements of all candidate profiles with those of the $\mathcal{D}$-optimal conjoint designs in Table 1 revealed that the values are identical for the same sample size and level setting. As a result, the $\mathcal{D}$-optimal arrangements of all candidate profiles are statistically as efficient as the designs in Table 1. The reason for this design equivalence lies in the fact that the $\mathcal{D}$-optimal conjoint designs are constructed for the random respondent effects model (2) embracing main effects only. It is therefore also most likely that there are still other conjoint designs equivalent to the $\mathcal{D}$-optimal ones. The $\mathcal{D}$-optimal arrangements of all candidate profiles are displayed in Tables 7 and 8 for the homogeneous and heterogeneous level setting, respectively. In Wu and Hamada (2000), page 253, the 81 run $\mathcal{D}$-optimal 3^4 block design containing 27 blocks of size three is also listed.

Table 7: $\mathcal{D}$-optimal arrangement of the 81 candidate profiles in the homogeneous level setting.

	Attributes					Attributes					Attributes			
Block	1	2	3	4	Block	1	2	3	4	Block	1	2	3	4
1	1	1	1	1	10	1	1	2	1	19	1	1	3	1
1	3	2	2	2	10	3	2	3	2	19	3	2	1	2
1	2	3	3	3	10	2	3	1	3	19	2	3	2	3
2	2	1	1	1	11	2	1	2	1	20	2	1	3	1
2	1	2	2	2	11	1	2	3	2	20	1	2	1	2
2	3	3	3	3	11	3	3	1	3	20	3	3	2	3
3	3	1	1	1	12	3	1	2	1	21	3	1	3	1
3	2	2	2	2	12	2	2	3	2	21	2	2	1	2
3	1	3	3	3	12	1	3	1	3	21	1	3	2	3
4	1	2	1	1	13	1	2	2	1	22	1	2	3	1
4	3	3	2	2	13	3	3	3	2	22	3	3	1	2
4	2	1	3	3	13	2	1	1	3	22	2	1	2	3
5	2	2	1	1	14	2	2	2	1	23	2	2	3	1
5	1	3	2	2	14	1	3	3	2	23	3	1	2	3
5	3	1	3	3	14	3	1	1	3	23	1	3	1	2
6	3	2	1	1	15	3	2	2	1	24	3	2	3	1
6	2	3	2	2	15	2	3	3	2	24	2	3	1	2
6	1	1	3	3	15	1	1	1	3	24	1	1	2	3
7	1	3	1	1	16	1	3	2	1	25	1	3	3	1
7	3	1	2	2	16	3	1	3	2	25	3	1	1	2
7	2	2	3	3	16	2	2	1	3	25	2	2	2	3
8	2	3	1	1	17	2	3	2	1	26	2	3	3	1
8	1	1	2	2	17	1	1	3	2	26	1	1	1	2
8	3	2	3	3	17	3	2	1	3	26	3	2	2	3
9	3	3	1	1	18	3	3	2	1	27	3	3	3	1
9	2	1	2	2	18	1	2	1	3	27	2	1	1	2
9	1	2	3	3	18	2	1	3	2	27	1	2	2	3

Table 8: $\mathcal{D}$-optimal arrangement of the 72 candidate profiles in the heterogeneous level setting.

	Attributes					Attributes					Attributes			
Block	1	2	3	4	Block	1	2	3	4	Block	1	2	3	4
1	1	1	1	3	9	1	3	1	3	17	1	1	3	1
1	2	2	3	2	9	2	1	2	1	17	1	3	4	2
1	2	3	2	1	9	2	2	4	2	17	2	2	1	3
2	1	1	2	3	10	1	3	2	3	18	1	1	4	1
2	2	2	4	1	10	2	1	1	2	18	1	3	3	2
2	2	3	1	2	10	2	2	3	1	18	2	2	2	3
3	1	1	3	3	11	1	3	3	3	19	1	1	1	2
3	2	2	1	1	11	2	1	4	1	19	1	3	2	1
3	2	3	4	2	11	2	2	2	2	19	2	2	3	3
4	1	1	4	3	12	1	3	4	3	20	1	1	2	1
4	2	2	2	1	12	2	1	3	1	20	1	3	1	2
4	2	3	3	2	12	2	2	1	2	20	2	2	4	3
5	1	2	1	3	13	1	2	4	1	21	1	1	2	2
5	2	1	4	2	13	1	3	2	2	21	1	2	3	1
5	2	3	3	1	13	2	1	1	3	21	2	3	1	3
6	1	2	2	3	14	1	2	3	2	22	1	1	1	1
6	2	1	3	2	14	1	3	1	1	22	1	2	4	2
6	2	3	4	1	14	2	1	2	3	22	2	3	2	3
7	1	2	3	3	15	1	2	2	2	23	1	1	4	2
7	2	1	2	2	15	1	3	4	1	23	1	2	1	1
7	2	3	1	1	15	2	1	3	3	23	2	3	3	3
8	1	2	4	3	16	1	2	1	2	24	1	1	3	2
8	2	1	1	1	16	1	3	3	1	24	1	2	2	1
8	2	3	2	2	16	2	1	4	3	24	2	3	4	3

6 Summary and conclusion

In this paper, we constructed $\mathcal{D}$-optimal designs for rating-based conjoint experiments by means of the linear random respondent effects model. In this model, a respondent effect is included to represent the fact that profile ratings from the same respondent are correlated. The resulting $\mathcal{D}$-optimal conjoint designs are statistically more efficient than $\mathcal{D}$-optimal completely randomized designs, which are based on the linear model without a respondent effect. Focussing on the $\mathcal{D}$-optimal conjoint designs, we could observe how many respondents are optimally needed for a specific conjoint setting and which and how many profiles are administered to each of them. Generally, it turns out that it is statistically most efficient to administer three profiles per respondent provided the sample size n is not too small ($n > 10$). If the sample size is not a multiple of three, four profiles can be assigned to one or two respondents. Consequently, if there are respondent costs, instead of determining the number of observations to carry out, it is more sensible to determine the number of respondents to appoint and to assign to each of them three profiles. Furthermore, to select the optimal profiles for quite large conjoint experiments ($n > 40$), it is advisable to replicate a smaller $\mathcal{D}$-optimal design ($n > 10$) to save on computation time. Finally, trying to save more on computation time by constructing $\mathcal{D}$-optimal completely randomized designs and arbitrary distributing three profiles per respondent, or sometimes four profiles in case the sample size is not a multiple of three, is on average statistically very inefficient.

References

Atkinson, A. C. and Donev, A. N. (1992). *Optimum Experimental Designs*, Oxford U.K.: Clarendon Press.

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

Cochran, W. G. and Cox, G. M. (1957). *Experimental designs*, New York: Wiley.

Entholzner, M., Benda N., Schmelter T. and Schwabe, R. (2004). A note on designs for estimating population parameters, working paper, Otto von Guericke University, Institute for Mathematical Stochastics, Magdeburg, Germany.

Gilmour, S. G. and Trinca, L. A. (2000). Some practical advice on polynomial regression analysis from blocked response surface designs, *Communications in Statistics: Theory and Methods* **29**: 2157–2180.

Goos, P. (2002). *The Optimal Design of Blocked and Split-plot Experiments*, New York: Springer.

Goos, P. and Vandebroek, M. (2001a). D-optimal response surface designs in the presence of random block effects, *Computational Statistics and Data Analysis* **37**: 433–453.

Goos, P. and Vandebroek, M. (2001b). Optimal split-plot designs, *Journal of Quality Technology* **33**: 436–450.

Goos, P. and Vandebroek, M. (2004). Outperforming completely randomized designs, *Journal of Quality Technology* **36**: 12–26.

Green, P. E., Krieger, A. M. and Wind, Y. (2001). Thirty years of conjoint analysis: reflections and prospects, *Interfaces* **31**: 56–73.

Gustafsson, A., Herrmann A. and Huber, F. (2003). Conjoint analysis as an instrument of market research practice, in *Conjoint Measurement: Methods and Applications,* 3rd edition, edited by Gustafsson, A., Herrmann, A. and Huber, F., Berlin: Springer, pp. 5–46.

Huber, F., Herrmann A. and Gustafsson, A. (2003). On the influence of the evaluation methods in conjoint design – Some empirical results, in *Conjoint Measurement: Methods and Applications,* 3rd edition, edited by Gustafsson, A., Herrmann, A. and Huber, F., Berlin: Springer, pp. 209–234.

Lenk, P. J., DeSarbo, W. S., Green, P. E. and Young, M. R. (1996). Hierarchical Bayes conjoint analysis: recovery of partworth heterogeneity from reduced experimental designs, *Marketing Science* **15**: 173–191.

Liski, E. P., Mandal, N. K., Shah, K. R. and Sinha, B. K. (2002). Topics in optimal design, *Lecture Notes in Statistics* **163**, New York: Springer.

Malhotra, N. K. (1986). An approach to the measurement of consumer preferences using limited information, *Journal of Marketing Research* **23**: 33–49.

Wittink, D. R. and Cattin, P. (1989). Commercial use of conjoint analysis: an update, *Journal of Marketing* **53**: 91–96.

Wittink, D. R., Vriens, M. and Burhenne, W. (1994). Commercial use of conjoint analysis in Europe: results and critical reflections, *International Journal of Research in Marketing* **11**: 41–52.

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