

Serial choice conjoint analysis for estimating discrete choice models

Michiel C.J. Bliemer

Delft University of Technology, m.c.j.bliemer@tudelft.nl
The University of Sydney, m.bliemer@itls.usyd.edu.au
Goudappel Coffeng BV, mbliemer@goudappel.nl

John M. Rose

The University of Sydney, j.rose@itls.usyd.edu.au

Abstract – Stated choice experiments can be used to estimate the parameters in discrete choice models by showing hypothetical choice situations to respondents. These attribute levels in each choice situation are determined by an underlying experimental design. Often, an orthogonal design is used, although recent studies have shown that better experimental designs exist, such as efficient designs. These designs provide more reliable parameter estimates. However, they require prior information about the parameter values which is often not readily available. Serial efficient designs are proposed in this paper in which the design is updated during the survey. In contrast to adaptive conjoint, serial conjoint only changes the design across respondents, not within-respondent thereby avoiding endogeneity bias as much as possible. After each respondent, new parameters are estimated and used as priors for generating a new efficient design. Results using the multinomial logit model show that using such a serial design, using zero initial prior values, provides the same reliability of the parameter estimates as the best efficient design (based on the true parameters). Any possible bias can be avoided by using an orthogonal design for the first few respondents. Serial designs do not suffer from misspecification of the priors as they are continuously updated. The disadvantage is the extra implementation cost of an automated parameter estimation and design generation procedure in the survey. Also, the respondents have to be surveyed in a mostly serial fashion instead of all parallel.

Keywords: stated choice experiment, discrete choice models, orthogonal design, efficient design, serial design

Introduction

Discrete choice models need no introduction given the vast amount of literature on the topic (e.g., Ben-Akiva and Lerman, 1985; Hensher *et al.*, 2005; Louviere *et al.*, 2000; Train, 2003) and numerous applications of such models in a wide range of fields, such as marketing, transportation, health economics, etc. While estimation of discrete choice models has received much attention (mainly in the transportation literature), only in the last decade has procedures for obtaining data specific for estimating such models become an increasingly popular field of study (mainly in the marketing literature). Aside from revealed preference data, where choices are observed in real markets, there has been a growing interest in stated preference data, with constructed choice situations in a hypothetical market. Generating these choice situations in stated choice (SC) experiments requires knowledge in the field of experimental design. Typically, each respondent in a SC experiment is faced with multiple choice situations where the respondent is asked to make a choice between two or more alternatives, described by attributes and their levels. The levels for all attributes and alternatives typically are drawn from an underlying experimental design that has to be generated beforehand by the analyst.

Several researchers have focused on generating experimental designs for SC experiments. Traditionally, orthogonal designs have been used (see e.g., Louviere *et al.*, 2000), which aim to minimize the correlation between the levels of any two attributes. Historically, orthogonal designs were generated for linear models however since discrete choice models are not linear, researchers have shown that they are not the most suitable design to use for SC studies. This fact has been acknowledged by, amongst others, Bunch *et al.* (1994), Kuhfeld *et al.* (1994) and Huber and Zwerina (1996) who focused on the efficiency of the design by linking it explicitly to the discrete choice model to be estimated. These researchers defined efficiency in terms of the reliability of the parameter estimates; roughly speaking the aim is to minimize the standard errors of the parameter estimates. Since then many researchers have focused on the generation of efficient designs, including as Bliemer and Rose (2008), Bliemer *et al.* (2009), Carlsson and Martinsson (2003), Ferrini and Scarpa (2006), Johnson *et al.* (2006), Kanninen (2002, 2005), Kessels *et al.* (2006, *in press*), Rose and Bliemer (2008), Rose *et al.* (2008), Sándor and Wedel (2001, 2002, 2005), Toner *et al.* (1999), and Yu *et al.* (2009). Original research focused only on designs for the multinomial logit (MNL) model, but recent publications have seen a shift towards mixed multinomial logit (MMNL) models and to a lesser extent nested logit (NL) models. Other researchers do not focus on optimal attribute levels, but rather focus on the optimal choice probabilities for each alternative (e.g., Johnson *et al.*, 2006; Kanninen, 2002, 2005; and Toner *et al.*, 1999).

In order to generate efficient designs, prior parameter values (best guesses) are needed. Early work (e.g., Huber and Zwerina, 1996) assumed fixed or precisely known values for the parameters. Later, Sándor and Wedel (2001) introduced random prior distributions to indicate uncertainty about these priors leading to so-called Bayesian efficient designs. Still, one has to decide which parameter distributions are to be used, with pilot studies typically needed to determine good prior distributions.

All of these efficient designs let go of the principle of orthogonality. Nevertheless, a separate stream of research on experimental design has recently emerged by Burgess and Street (2003), Street and Burgess (2004), and Street *et al.* (2001, 2005) which seeks to locate optimal orthogonal designs without requiring any prior information (basically assuming prior values equal to zero). Whilst this approach has been adopted frequently in the marketing and

health economics literature, it does have some drawbacks. For example, it is only suitable for unlabeled experiments with generic parameters, and the designs can be shown to lose efficiency (as defined by the wider literature on experimental designs for SC studies) if the priors are not equal to zero in practice (which is essentially always the case). Still, the property that no priors are required is appealing to many, as these priors are usually not readily available.

In order to overcome the problem of having to specify *a priori* prior parameters with more or less uncertainty but still some level of statistical efficiency in generated designs, *adaptive designs* (mostly referred to as adaptive conjoint analysis) have been introduced, see e.g., Johnson (1987, 1991), and Green *et al.* (1991). The choice situations in these designs are not fixed, but are adapted based on answers given by the respondent in earlier choice situations. However, such design methods are not without limitations with Toubia *et al.* (2003) stating:

“To understand this result we first recognize that any adaptive question-design method is potentially subject to endogeneity bias. Specifically, the q th question depends upon the answers to the first $q - 1$ questions. This means that the q th question depends, in part, on any response errors in the first $q - 1$ questions. This is a classical problem, which often leads to bias (see, for example, Judge *et al.* 1985, p. 571). Thus, adaptivity represents a tradeoff: We get better estimates more quickly, but with the risk of endogeneity bias.”

Adaptive conjoint can be distinguished by changes in the questions asked both within- and across respondents over the course of the study. In order to prevent endogeneity bias as much as possible, we propose the use of an adaptive process in which the design is only changed across respondents, which we term *serial efficient designs*. Such a procedure has been suggested by others, but has not been explored in more detail. Unlike other typical designs such as orthogonal and regular (Bayesian) efficient designs, a serial efficient design is no longer the same for each respondent. At the beginning of the survey, no information is assumed. An orthogonal design for example be used for the first respondent, or alternatively an efficient design based on zero-valued priors. After completion of the S choice situations by this first respondent, the parameters are estimated based on his or her observed choices. Parameters that turn out to be statistically significant are then used as priors in determining the next design whilst parameters that are not statistically significant are assumed to be zero. Based on these new priors, a new efficient design can be generated and given to the next respondent. The data from each additional respondent is then pooled with the data from previously surveyed respondents and new models estimated, after which a new design is generated and given to the next surveyed respondent. There may still be some risk of endogeneity bias in such a procedure, however, by requiring attribute level balance¹ for each new design, the choice situations will still cover a wide range of the utility space and will not concentrate (as in adaptive conjoint) on certain parts of utility space, thus keeping any possible bias to a minimum. Hauser and Toubia (2005) state that endogeneity bias in general exists in adaptive conjoint analysis, but also mention that adaptive utility-balanced choice questions do not appear to be biased in this way.

The paper is outlined as follows. First, the MNL model will be discussed, mainly to introduce the require notation. The next section will briefly state the parameter estimation process, which is needed to update the priors for constructing the serial designs. Then, the generation of efficient designs is discussed and the algorithm for the serial efficient design procedure is

¹ Attribute level balance means that each level for each attribute appears an equal number of times over all choice situations, hence a respondent does not see for example mainly high or mainly low attribute levels.

stated in more detail. Two case studies are presented illustrating the potential of serial designs and compare them to using (optimal) orthogonal and regular fixed efficient designs for creating stated choice data and parameter estimation. Finally, the results are summarized and discussed with some advantages and disadvantages of the serial design approach.

The multinomial logit model

Consider N respondents that have to make a choice between J alternatives in S different choice situations. Let u_{nsj} denote the perceived utility of respondent n for alternative j in choice situation s . Furthermore, assume that this perceived utility consists of a systematic utility component, v_{nsj} , and a random component, ε_{nsj} ,

$$u_{nsj} = v_{nsj} + \varepsilon_{nsj}. \quad (1)$$

The systematic utility is often assumed to be composed of a linear combination of attribute values and associated weights,

$$v_{nsj} = x_{nsj}'\beta, \quad (2)$$

where $x_{nsj} \in R^K$ is a vector of K attribute values, and $\beta \in R^K$ is a vector of weights (parameters). Each respondent is assumed to choose the alternative that maximizes his or her perceived utility. The probability of a respondent choosing a certain alternative depends on the assumptions made on the random components ε_{nsj} . Under the strict assumption that these random components are independently and identically extreme value type I distributed, the probability p_{nsj} that respondent n chooses alternative j in choice situation s can be written as (see McFadden, 1974):

$$p_{nsj} = \frac{\exp(v_{nsj})}{\sum_{i=1}^J \exp(v_{nsi})}. \quad (3)$$

More sophisticated logit models exist, e.g., generalized extreme value models or MMNL models, however for simplicity the MNL model will be used in this paper. Hereby we ignore the repeating nature of successive choice observations by a single respondent in a SC experiment. The panel MMNL model is suitable to take there correlations between choice observations into account, however, this model is much more complex and optimal experimental design is computationally much more expensive (see Bliemer and Rose, 2008). Further, Bliemer and Rose (2008) and Rose et al. have shown that designs generated for the MNL model offers similar efficiency levels to those generated specifically for MMNL models. Nevertheless, for a real-life case study the use of MMNL designs would probably not represent a significant issue, however for our analysis we adopt a Monte Carlo approach, repeating the process 100 times, making the computation times with the panel MMNL model prohibitively large. Therefore, we restrict ourselves to the simpler MNL model and note that there is no reason to believe that results found in this paper are not transferrable to more sophisticated model types.

Model estimation

In order to estimate the unknown parameters β in the MNL model formulated above, a SC experiment can be conducted. In such an experiment, the attribute levels x_{nsj} are input from

an underlying experimental design and presented to the respondents in a survey. The respondents then state their preferred alternative in each choice situation. Let $y_n \in R^{SJ}$ denote the vector of choices made by respondent n in a sequence of S choice situations, where $y_{nsj} = 1$ if the respondent chooses alternative j in choice situation s , and zero otherwise. These choice outcomes can be used to estimate parameters β . More specifically, the parameters can be estimated by maximizing the following loglikelihood function:

$$\ell_N(\beta | X_N, Y_N) = Y_N' \log P_N(X_N | \beta), \quad (4)$$

where² $X_N = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}$, $Y_N = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix}$, and $P_N = \begin{pmatrix} p_1 \\ \vdots \\ p_N \end{pmatrix}$, with $p_n \in R^{SJ}$ the MNL probabilities for respondent n from Eqn. (3), and $x_n \in R^{(SJ) \times K}$ the attribute levels for each respondent n ,

$$x_n = \begin{pmatrix} x_{n111} & x_{n112} & \cdots & x_{n11K} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1J1} & x_{n1J2} & \cdots & x_{n1JK} \\ x_{n211} & x_{n212} & \cdots & x_{n21K} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n2J1} & x_{n2J2} & \cdots & x_{n2JK} \\ \vdots & \vdots & \vdots & \vdots \\ x_{nSJ1} & x_{nSJ2} & \cdots & x_{nSJK} \end{pmatrix}. \quad (5)$$

This matrix of attribute levels can also be referred to as the experimental design for respondent n .

Since the function in Eqn. (4) is concave (see e.g., Train, 2003), in order to find the maximum the Newton-Rhaphson iterative procedure can be applied by using the first and second derivatives. The vector of first derivatives, the gradient $g_N \in R^K$, and the matrix of second derivatives, the Hessian $h_N \in R^{K \times K}$, are given by

$$g_N(\beta | X_N, Y_N) = \frac{\partial \ell_N}{\partial \beta} = X_N' (Y_N - P_N), \quad \text{and} \quad (6)$$

$$h_N(\beta | X_N) = \frac{\partial^2 \ell_N}{\partial \beta \partial \beta'} = -Z_N' \text{diag}(P_N) Z_N, \quad (7)$$

where $Z_N = \begin{pmatrix} z_1 \\ \vdots \\ z_N \end{pmatrix}$, with $z_{nsjk} = x_{nsjk} - \sum_{i=1}^J p_{nsi} x_{nsik}$.

² Upper case variables in this paper denote accumulated data, e.g., Y_n represents all data from y_1 to y_n .

Experimental design

So far, the experimental design was assumed given. The question now is, which designs x_n , $n = 1, \dots, N$, will yield the most accurate and/or reliable parameter estimates? The design that yields reliable parameter estimates (i.e., small standard errors) for a fixed sample size is called an efficient design. Traditionally, the same design is used for each respondent and therefore only a single design is constructed. The most widely used design type is a so-called *orthogonal design* which aims to minimize the correlation between the levels of two distinct attributes. This type of design is optimal (most efficient) for estimating linear models, however the MNL model is nonlinear and therefore this design will typically not be the most efficient design possible. More efficient designs can be found by taking the model type explicitly into account. These *efficient designs* are usually found by minimizing the so-called D-error, leading to a D-efficient design. The D-error can be computed by taking the determinant of the asymptotic variance-covariance (AVC) matrix for a single respondent n , $\Omega_n \in R^{K \times K}$, and normalizing this value to the number of parameters, K ,

$$d(x_n | \beta) = [\det(\Omega_n)]^{1/K} = \left(\frac{1}{\det(-h_n(\beta | x_n))} \right)^{1/K}. \quad (8)$$

As stated in the equation above, the AVC matrix is the inverse of the Fisher information matrix, which is the negative Hessian matrix. It is important to observe that h_n does not depend on y_n , therefore the D-error can be determined without conducting any surveys.³ In other words, the efficiency of the design can be evaluated beforehand. However, h_n can only be determined if the parameter values, β , are known. Since the aim is to estimate these, these values are clearly unknown. Hence, prior values $\tilde{\beta}$ have to be assumed as best guesses. These priors can be obtained from literature or pilot studies. Lack of knowledge about these priors is one of the aspects that limits the generation of efficient designs.

In Figure 1 the procedures for estimating parameters using different types of experimental designs are depicted. When using an efficient or serial design priors are needed, but for the serial design one can start with zero priors, while for the efficient design more effort is needed to determine good priors. Orthogonal designs do not require any priors (step 1 is missing). The serial design is updated for each respondent (using updated priors), while this updating step 4 is missing when using an efficient or orthogonal design. In the last step, the parameters are estimated, together with their standard errors. As mentioned before, in general the lower these standard errors, the more efficient the design.

In the next section, the procedure for determining serial designs will be discussed in detail.

³ It has been shown in Bliemer et al. (2009) and Sándor and Wedel (2002) that both the nested logit model and the cross-sectional mixed logit model share this same property. However, as discussed in Bliemer and Rose (2008), in the panel mixed logit model the dependency on y_n remains, such that sampling is required.

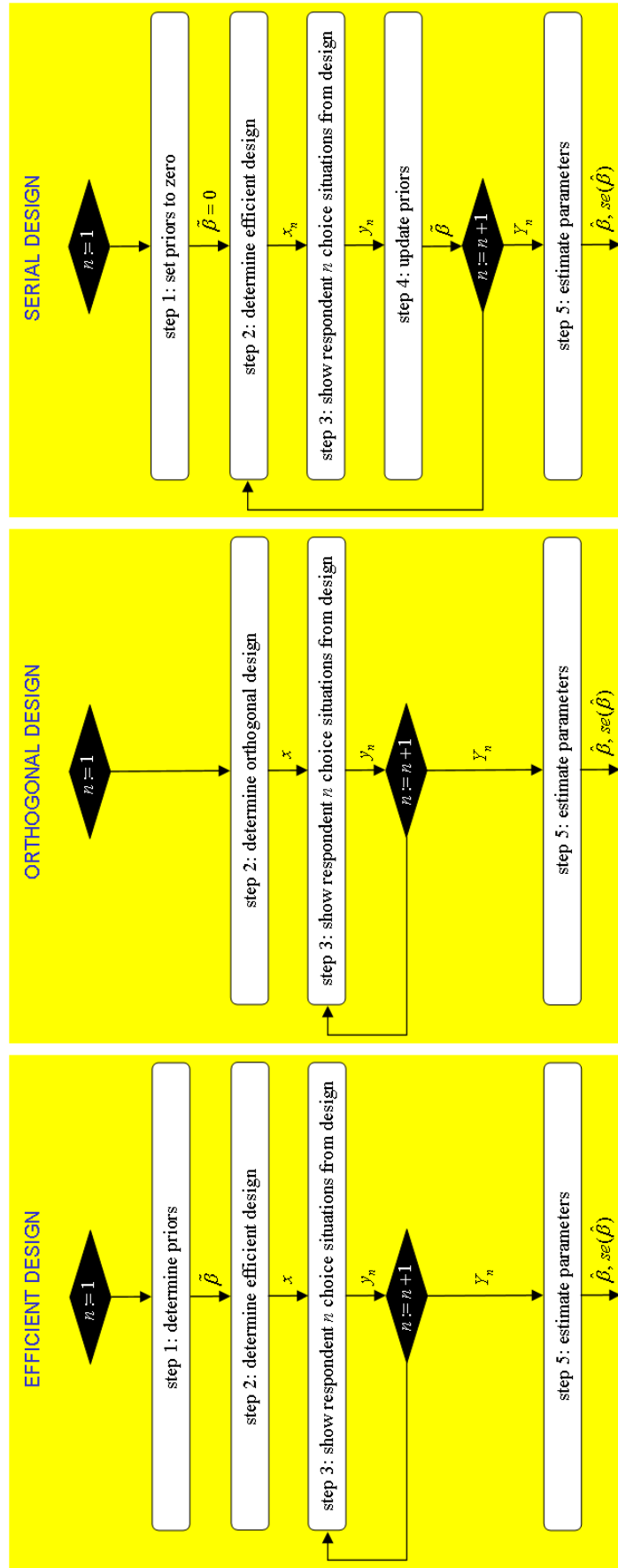


Figure 1: Procedure for estimating parameters using an efficient, orthogonal, or serial design

Generating serial efficient designs

The five steps in Figure 1 will now be described in more detail.

Step 0: Initialization

Specify the utility function in Eqn. (2) and determine the dimensions of the design, i.e., the number of choice situations, S , and the possible levels of each attribute. Let Λ denote the set of feasible designs, constrained by the specified attribute levels and attribute level balance. Set $n = 1$. Let N be the sample size. Set $X_0 = Y_0 = P_0 = Z_0 = \emptyset$.

Step 1: Initialize priors

Assuming no prior information on parameter values, set $\tilde{\beta} = 0$.

Step 2: Determine efficient design

Determine design x_n such that the D-error in Eqn. (8) is minimized for the given $\tilde{\beta}$. Finding an optimal design is very difficult, therefore we settle for the most efficient design that we can find after a few thousand design evaluations. These designs are not chosen randomly from the set of feasible designs, Λ , but are optimized using local search techniques. Swapping algorithms that switch two attribute levels for the same attribute are easy to implement, maintain attribute level balance, and find relatively efficient designs rather quick. The design for the previous respondent, x_{n-1} , is included in the search space as a good starting point for swapping.

Pool all the design data, $X_n = \begin{pmatrix} X_{n-1} \\ x_n \end{pmatrix}$.

Step 3: Observe choices from respondent

Present choice situations dictated by design x_n to respondent n and collect his or her choice data, y_n .

Pool all the observed choices, $Y_n = \begin{pmatrix} Y_{n-1} \\ y_n \end{pmatrix}$.

Step 4: Update priors

Estimate parameters based on the design, x_n , and the choice data, y_n , using the Newton-Raphson iterative procedure:

- (a) Set $\beta = \tilde{\beta}$ as an initial solution.
- (b) Compute the probabilities corresponding to these parameter values using Eqn. (3).
- (c) Compute $z_{nsjk} = x_{nsjk} - \sum_{i=1}^J p_{nsi} x_{nsik}$.
- (d) Pool the variables: $Z_n = \begin{pmatrix} Z_{n-1} \\ z_n \end{pmatrix}$, $P_n = \begin{pmatrix} P_{n-1} \\ P_n \end{pmatrix}$.
- (e) Compute gradient: $g_n = X_n' (Y_n - P_n)$.
- (f) Compute Hessian: $h_n = -Z_n' \text{diag}(P_n) Z_n$.
- (g) Determine asymptotic variance-covariance matrix: $\Omega_n = -h_n^{-1}$.
- (h) Update parameter estimate: $\beta := \beta + \Omega_n g_n$.
- (i) Convergence: if $\|g_n\| < \delta_1$, then set $\tilde{\beta} = \beta$ and continue to Step 4(j), otherwise return to Step 4(b).

- (j) Let $se(\beta_k)$ be the standard error for parameter k , i.e. the square root of the k -th diagonal element of matrix Ω . If $|\beta_k / se(\beta_k)| < 1.96$, then β_k is not statistically significant, such that we set the prior to zero, $\tilde{\beta}_k = 0$.
- (k) If $n < N$, then set $n := n + 1$ and return to Step 2. Otherwise, continue to Step 5.

Step 5: Estimate parameters

The parameter estimates are readily available from the previous step; the estimates are $\hat{\beta} = \beta$.

The above procedure describes how to apply serial efficient designs in practice. For the results in this paper, we have not conducted any real surveys, however we rely on simulation results. Step 3 in the above procedure therefore is replaced by a sample generator that will simulate choices of respondents, assuming they behave as described by the MNL model. Below is the description of the rather straightforward sample generator, where $\bar{\beta}$ is the assumed to be the true set of parameters for the population.

Step 3: Sample generator for respondent n

- (a) Compute systematic utilities: $v_{nsj} = x_{nsj}' \bar{\beta}$.
- (b) Compute error components: $\varepsilon_{nsj} = -\log(-\log(\omega_{nsj}))$, where ω_{nsj} is a (uniformly distributed) random number between 0 and 1.
- (c) Compute utilities perceived by respondent n : $u_{nsj} = v_{nsj} + \varepsilon_{nsj}$.
- (d) For each s , determine the alternative with the highest utility: $j_{ns}^* = \arg \max_j (u_{nsj})$. In case of multiple alternatives with highest utility, a random selection is made.
- (e) Define $y_{nsj} = \begin{cases} 1, & \text{if } j = j_{ns}^* \\ 0, & \text{otherwise.} \end{cases}$

Case studies

In this section two examples are shown. In the first example a labeled experiment with some alternative-specific parameters is considered while in the second example an unlabeled experiment with all generic parameters is assumed. In both examples, an orthogonal design is constructed with D-efficient designs constructed based on priors that correspond to the true parameters. This latter set of designs is expected to offer better performance. In each example, serial efficient designs are constructed assuming 80 respondents. Since the sample generator contains a random error component, the generated choices will be different for each time we generate a sample, leading to different priors each time, and also different serial designs. Therefore, a Monte Carlo approach is adopted in which we repeat the whole process 100 times. The outcomes in terms of parameter accuracy (the actual values) and parameter reliability (in terms of t -values) will be discussed.

Example 1

Consider the following utility functions with eight attributes and seven (generic and alternative-specific) parameters to estimate:

$$\begin{aligned}
 v_{ns1} &= \beta_1 + \beta_3 x_{ns1,3} + \beta_4 x_{ns1,4} + \beta_5 x_{ns1,5} \\
 v_{ns2} &= \beta_2 + \beta_3 x_{ns2,3} + \beta_4 x_{ns2,4} + \beta_6 x_{ns2,6} \\
 v_{ns3} &= \beta_3 x_{ns3,3} + \beta_7 x_{ns3,7}
 \end{aligned}$$

The following attribute levels are assumed:

$$x_{ns1,3}, x_{ns2,3}, x_{ns3,3} \in \{6, 8, 10, 12\}, x_{ns1,4}, x_{ns2,4} \in \{4, 8\}, x_{ns1,5}, x_{ns2,6}, x_{ns3,7} \in \{0, 1\}.$$

The following values for the (true) parameters are assumed:

$$\bar{\beta} = (1.2, 0.8, -0.6, -0.4, 0.3, 0.8, -1.0)'$$

An attribute-level balanced orthogonal and a D-efficient design (assuming $\bar{\beta}$ as prior) with 12 choice situations have been generated using the Ngene software and are shown in Table 1 (in the same format as the matrix in Eqn. (5)). The D-error of the efficient design is 0.5147 which is significantly lower than the orthogonal design with a D-error of 0.8802. As usual, the orthogonal design is more efficient under the assumption of zero priors. In that case, the D-error would be 0.3572 compared to a D-error of 0.3965 for the efficient design. The designs are used for all respondents. Obviously, the serial efficient design is variable and therefore not listed here.

In each of the Monte Carlo runs, we simulate respondents and after each respondent the parameters are estimated again. In this way, we can observe how the parameter estimates converge to the true parameter values $\bar{\beta}$ as the sample size increases. As an example, the parameter estimates for $\hat{\beta}_2$ for 100 runs of up to 80 respondents are depicted in Figure 2, in which the horizontal line indicates the true parameter value. In general, the more respondents, the more accurate the parameter estimates become, although there exists quite some variability in the actual values obtained. Comparing the three different designs, it is clear that the orthogonal design has a larger spread of parameter estimates, while the efficient and serial designs seem to provide similar results.

In Figure 3 the average values of all seven parameter estimates are shown for each sample size and each design type. Particularly in the first four parameters there is some clear overestimation (in absolute value) in the serial design, which slowly disappears with larger sample sizes. With low sample sizes the priors are not of high quality yet and therefore the serial design may not be of high quality either. The result is that the information collected from the first respondents is not optimal. This may suggest using the orthogonal design for the first few respondents before estimating and updating any priors. Also, possible endogeneity bias may play a role although with higher sample sizes this bias appears to play no role. We will further investigate the overestimation or bias later in this section.

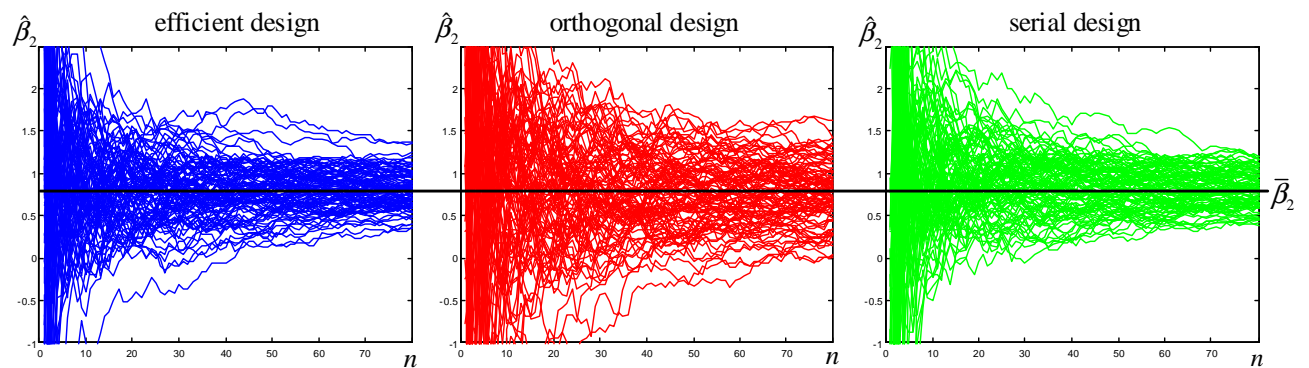


Figure 2: Parameter estimates of β_2 in Example 1 for different sample sizes (100 runs)

Table 1: Attribute levels of the orthogonal and efficient designs in Example 1

$s:$	$j:$	$k:$	orthogonal design							efficient design						
			1	2	3	4	5	6	7	1	2	3	4	5	6	7
1	1		1	0	6	8	0	0	0	1	0	8	4	1	0	0
1	2		0	1	6	8	0	0	0	0	1	8	8	0	1	0
1	3		0	0	6	0	0	0	1	0	0	6	0	0	0	1
2	1		1	0	10	4	1	0	0	1	0	10	8	1	0	0
2	2		0	1	10	8	0	1	0	0	1	10	4	0	0	0
2	3		0	0	8	0	0	0	1	0	0	12	0	0	0	0
3	1		1	0	10	8	1	0	0	1	0	12	4	1	0	0
3	2		0	1	8	4	0	1	0	0	1	6	8	0	0	0
3	3		0	0	10	0	0	0	1	0	0	8	0	0	0	1
4	1		1	0	12	4	0	0	0	1	0	8	8	0	0	0
4	2		0	1	6	4	0	0	0	0	1	10	4	0	1	0
4	3		0	0	12	0	0	0	1	0	0	10	0	0	0	1
5	1		1	0	6	4	0	0	0	1	0	10	4	0	0	0
5	2		0	1	8	8	0	1	0	0	1	6	8	0	1	0
5	3		0	0	12	0	0	0	0	0	0	8	0	0	0	0
6	1		1	0	8	4	0	0	0	1	0	6	4	0	0	0
6	2		0	1	12	4	0	1	0	0	1	6	8	0	1	0
6	3		0	0	6	0	0	0	1	0	0	6	0	0	0	0
7	1		1	0	8	8	1	0	0	1	0	8	8	1	0	0
7	2		0	1	6	4	0	1	0	0	1	12	4	0	0	0
7	3		0	0	8	0	0	0	0	0	0	12	0	0	0	0
8	1		1	0	10	8	0	0	0	1	0	12	4	0	0	0
8	2		0	1	12	4	0	0	0	0	1	8	8	0	0	0
8	3		0	0	8	0	0	0	0	0	0	12	0	0	0	1
9	1		1	0	8	8	1	0	0	1	0	12	8	1	0	0
9	2		0	1	12	8	0	0	0	0	1	12	4	0	1	0
9	3		0	0	12	0	0	0	1	0	0	10	0	0	0	1
10	1		1	0	12	4	1	0	0	1	0	10	4	1	0	0
10	2		0	1	8	8	0	0	0	0	1	8	8	0	1	0
10	3		0	0	6	0	0	0	0	0	0	8	0	0	0	0
11	1		1	0	12	8	0	0	0	1	0	6	8	0	0	0
11	2		0	1	10	8	0	1	0	0	1	12	4	0	0	0
11	3		0	0	10	0	0	0	0	0	0	6	0	0	0	1
12	1		1	0	6	4	1	0	0	1	0	6	8	0	0	0
12	2		0	1	10	4	0	0	0	0	1	10	4	0	0	0
12	3		0	0	10	0	0	0	0	0	0	10	0	0	0	0

The values in Figure 3 only show the average parameter estimates, however the variance of these parameter estimates plays an important role in assessing the reliability of the estimates. Figure 4 shows the absolute t -ratios for each parameter for different sample sizes, defined as the ratio of the true parameter (absolute) value and the standard error. The lower the standard error, the larger this ratio, and therefore the higher the reliability. The horizontal lines in the figure indicates a t -ratio of 1.96 as the 95%-level of statistical significance.

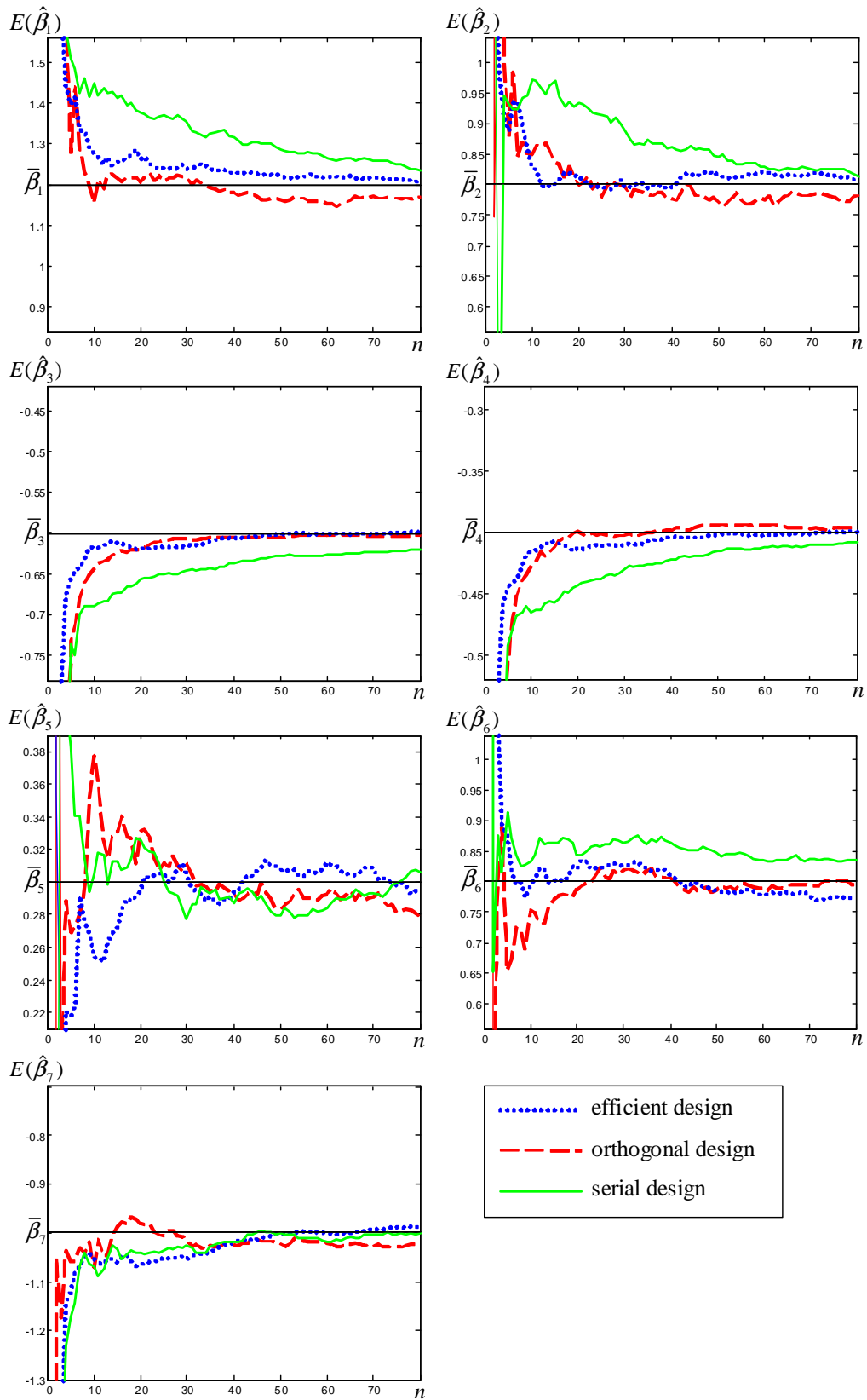


Figure 3: Average parameter estimates in Example 1 for different sample sizes

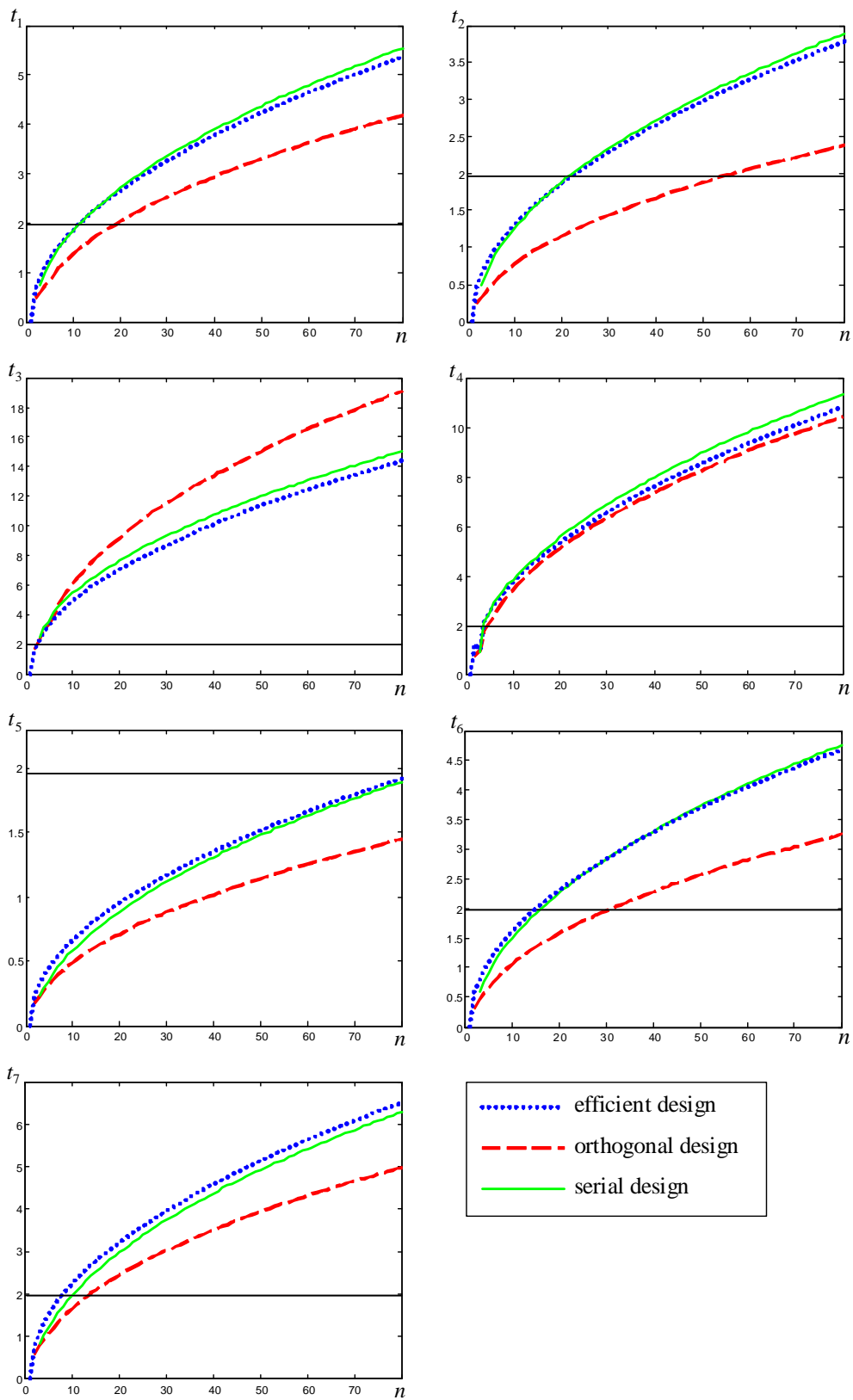


Figure 4: t -values in Example 1 for different sample sizes

The first observation that can be made is that the orthogonal design provides in general much less reliable parameter estimates, in line with Figure 2. Only for the third parameter does it perform better, but this parameter has already a very low standard error (the parameter is already statistically significant with a few respondents) and therefore is not of much concern. The second observation is that the performance of the efficient and serial designs are almost the same. The serial design, assuming no initial prior information, seems to catch up relatively quick with the efficient design.

Examining the points where the t -value exceeds 1.96, an indication of the required sample size can be obtained. These sample sizes are indicated in Figure 5. The fifth parameter is the most difficult parameter to estimate. The required sample sizes of the efficient and serial designs are almost identical, while the orthogonal design requires approximately twice as many respondents.

Coming back to the observed overestimation of the parameters with a serial design, another simulation is performed in which an orthogonal design is used for the first 20 respondents and then the design is updated with new priors from then on. This prevents bad and unreliable priors to be used in the generation of the serial design, which seems to cause the bias. The results are shown in Figure 6 for the second parameter, noting that all other parameters show similar results. The bias for the serial design has clearly disappeared. The efficiency in terms of t -ratios is low for the first 20 respondents, but quickly increases when the design becomes more efficient. This procedure seems to prevent bias while provides reliable parameter estimates.

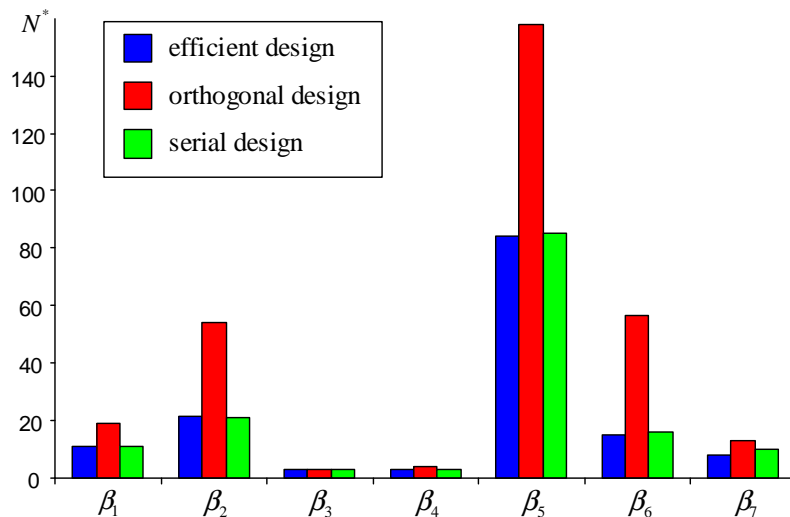


Figure 5: Sample sizes needed for parameters in Example 1 to be statistically significant

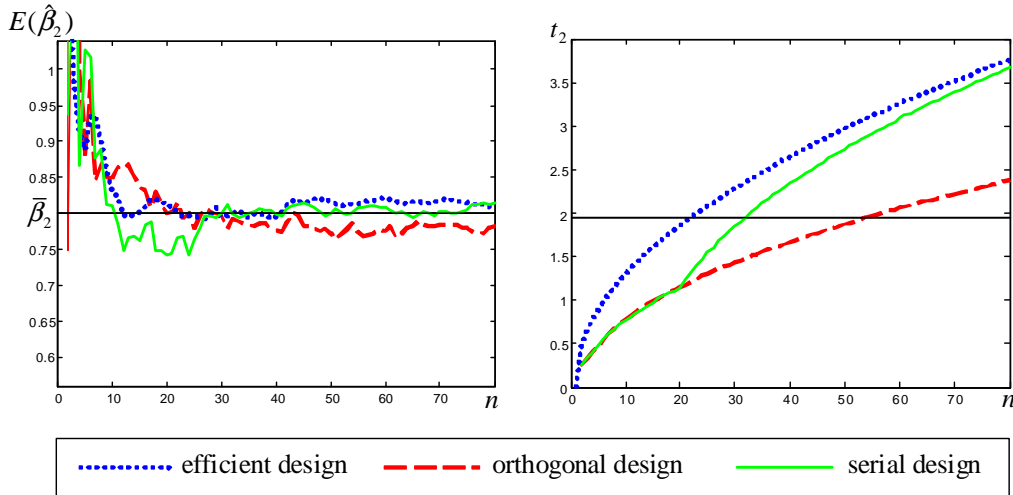


Figure 6: Serial design using an orthogonal design for the first 20 respondents

Example 2

Now consider the following two utility functions with four generic parameters:

$$v_{ns1} = \beta_1 x_{ns1,1} + \beta_2 x_{ns1,2} + \beta_3 x_{ns1,3} + \beta_4 x_{ns1,4}$$

$$v_{ns2} = \beta_1 x_{ns2,1} + \beta_2 x_{ns2,2} + \beta_3 x_{ns2,3} + \beta_4 x_{ns2,4}$$

All attributes are assumed to have three levels: 0, 1, or 2. The following values for the (true) parameters are assumed: $\bar{\beta} = (0.6, 0.4, -0.5, -0.2)'$.

For unlabelled experiments with all generic parameters, Street *et al.* (2005) describe a procedure to determine optimal orthogonal designs, although their definition of optimality is somewhat different from what we assume here as D-optimality. They define optimality in terms of maximizing the attribute level differences for the same attribute across alternatives, arguing that if an attribute level is the same across two or more alternatives, then no information is captured about trade-offs for that attribute. Such a definition does not require prior information when generating the design. Table 2 shows just such a design. Furthermore, a D-efficient design has also been generated and is shown in Table 2. The (optimal) orthogonal design has a D-error of 0.3143 (0.2222 if all zero priors are assumed) and the efficient design has a D-error of 0.2267 (0.1869 in case of all zero priors). As in the previous example, the average parameter estimates and the t -ratios are shown in Figures 7 and 8. The differences in results between designs are not as dramatic as in the previous example, most likely because the input was a ‘good’ orthogonal design, but the results are in line with the previous results. The orthogonal design performs worst, while the t -ratios of the serial design are again very close to the t -ratios of the efficient design. The required sample sizes are depicted in Figure 9.

Table 2: Attribute levels of the orthogonal and efficient designs in Example 2

$s:$	$j:$	$k:$	orthogonal design				efficient design			
			1	2	3	4	1	2	3	4
1	1		0	0	0	0	2	1	1	0
1	2		1	2	1	2	0	1	1	2
2	1		0	1	1	2	0	1	1	0
2	2		1	0	2	1	2	1	1	2
3	1		0	2	2	1	1	1	2	2
3	2		1	1	0	0	1	1	0	0
4	1		1	0	1	1	2	0	0	2
4	2		2	2	2	0	1	2	2	0
5	1		1	1	2	0	2	0	2	1
5	2		2	0	0	2	0	2	0	1
6	1		1	2	0	2	1	2	1	1
6	2		2	1	1	1	0	0	1	1
7	1		2	0	2	2	1	2	2	1
7	2		0	2	0	1	1	0	0	1
8	1		2	1	0	1	0	2	0	2
8	2		0	0	1	0	2	0	2	0
9	1		2	2	1	0	0	0	0	0
9	2		0	1	2	2	2	2	2	2

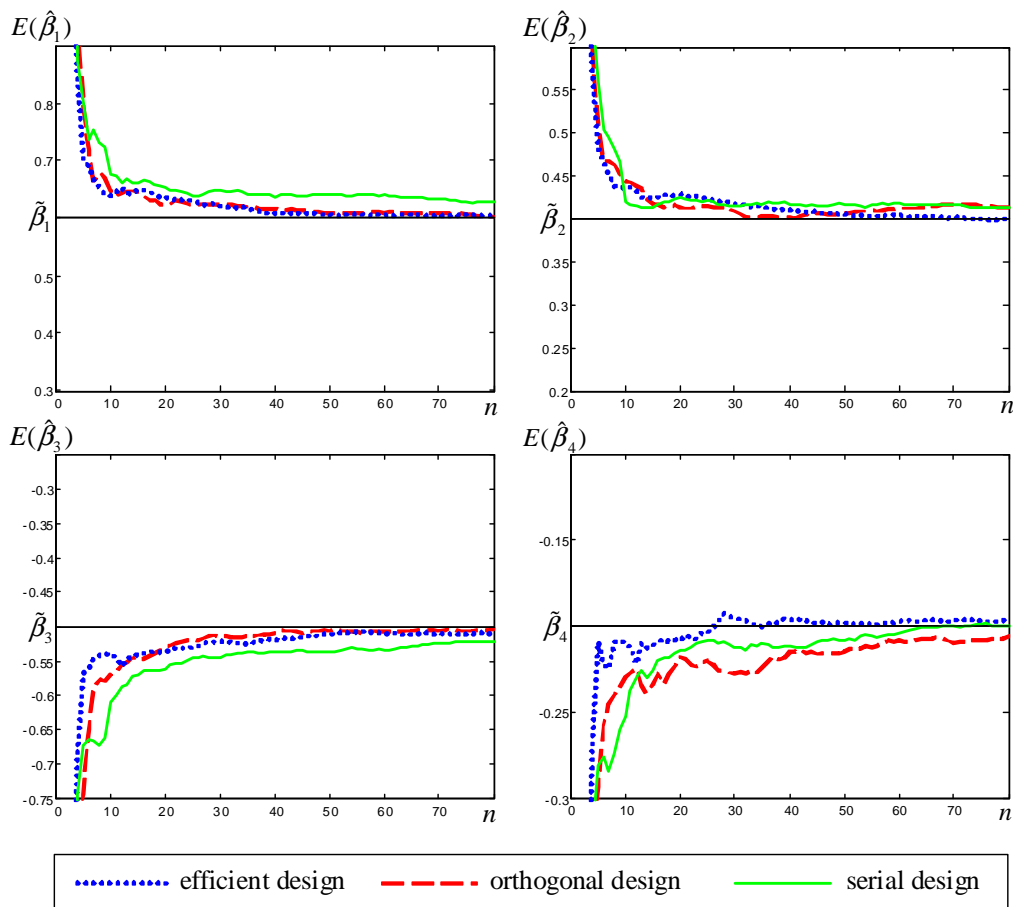


Figure 7: Average parameter estimates in Example 2 for different sample sizes

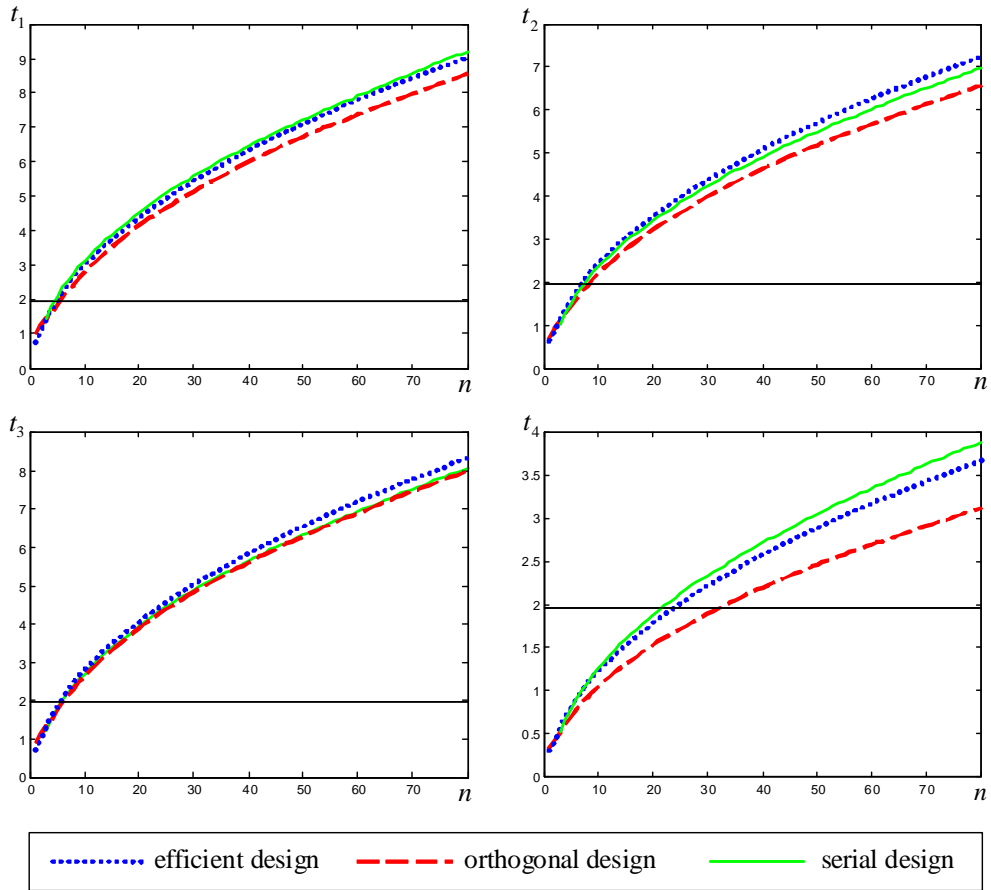


Figure 8: t -values in Example 2 for different sample sizes

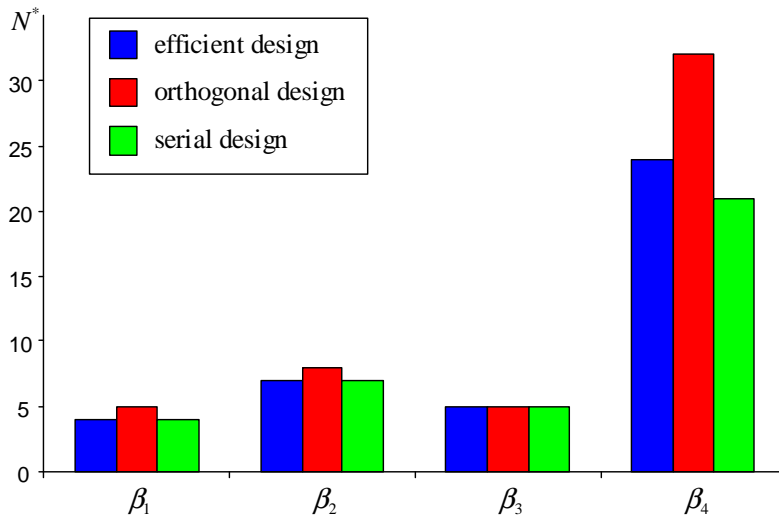


Figure 9: Sample sizes needed for parameters in Example 2 to be statistically significant

Summary and discussion

Orthogonal designs are still used most in practice, although these designs are clearly not the most efficient as has been shown several times in recent literature and again in this paper. However, constructing efficient experimental designs requires prior parameter information. Often, these priors are not readily available, and not all practitioners want to go through the process of first doing a pilot study in order to obtain some priors. Therefore, in this paper we have examined an automated updating procedure in which efficient designs are generated based on current information of the priors, i.e., parameters estimated from choice observations of previous respondents. No initial priors need be known, while in the end an efficient design (and data set) will result. The results show that a serial design is as good as an efficient design based on true parameter values. It is very unlikely that the true parameter values are known in advance and are at best only known with uncertainty. That serial design is able to produce results similar to the efficient design based on true parameters is therefore very promising. Serial designs are also not sensitive to misspecification of the priors, since the priors are updated continuously. Efficient designs lose efficiency when the priors turn out to be incorrect. Adopting a Bayesian approach in which the parameter priors are assumed to be random distributions, i.e., explicitly assuming that the priors are uncertain, would yield a design that is more robust against misspecification, but would also lose some efficiency. A serial Bayesian design seems an interesting direction for further research, in which not only the parameter estimates themselves are used to generate an efficient design, but also the standard error is used to indicate the level of uncertainty. For example, one can consider Bayesian priors following a normal distribution with means and standard deviations equal to the parameter estimates and the standard errors, respectively. However, generating a Bayesian efficient design is computationally intensive, as the Bayesian D-error has to be approximated with simulation. For more about computing the Bayesian efficiency of a design, see e.g. Bliemer *et al.* (2008).

The case studies also showed that there is a potential risk of overestimating the parameters (in the absolute sense), particularly at low sample sizes. This may be avoided by using an orthogonal design for the first few respondents, such that the first update of the priors is made based on sufficient data. In the case study, using an orthogonal design for the first 20 respondents before updating the design got rid of the overestimation problem completely, while still keeping most of the efficiency in parameter estimation.

The applicability is not restricted to only the MNL model, but also NL models and MMNL models could be estimated using serial experimental designs. For generating efficient designs for NL models, see Bliemer *et al.* (2009). For more on efficient designs for cross-sectional mixed logit models, see Sándor and Wedel (2002), and for the panel MMNL model, see Bliemer and Rose (2008). The MMNL models could potentially lead to computational problems, as in serial designs estimation of the parameters is done after each respondent. Estimation of MMNL models can be rather time consuming and may become impractical in serial design generation. Furthermore, generating efficient designs for MMNL models, particularly for panel MMNL models where sample generation is necessary (see Bliemer and Rose, 2008), is also computationally intensive. Depending on the size of the estimation problem, application of MMNL may be feasible or not (in terms of computation time). If there is no time restriction, for example if only a single respondent per day is surveyed, then clearly there is no problem.

Since orthogonal and efficient designs are kept fixed, the survey can be sent out to all respondents at the same time. These surveys can be done in parallel. Serial designs are generated based on choices from previous respondents, such that the survey cannot be sent out to all respondents at the same time. In case of a computer-aided personal interviewing (CAPI) technique, where an interviewer brings a laptop computer and visits respondents, this does not pose a problem. Internet surveys need a smarter implementation to avoid problems. For example, not after each respondent a new serial design is generated, but for example after every 10 respondents in an automated fashion, or every hour. The design does not change for respondents already filling out the survey, but only for newly logged in respondents.

Concluding, serial designs combine the benefits of orthogonal designs (no initial prior information required) and regular efficient designs (reliable parameter estimates). Also, a serial design is not sensitive to misspecification and will outperform Bayesian efficient designs that are only suboptimal. A disadvantage would be that more effort has to be put in the implementation of the survey (automated estimation and design generation) and that the serial character may prohibit large scale parallel surveying techniques. However, some form of serial design can always be applied, even by estimating the model parameters and updating the design manually at certain time intervals.

References

Bliemer, M.C.J., and Rose, J.M. (2008) Construction of experimental designs for mixed logit models allowing for correlation across choice observations. *Proceedings of the 87th Annual Meeting of the Transportation Research Board*, Washington, DC, USA.

Bliemer, M.C.J., Rose, J.M., and Hensher, D.A. (2009) Efficient stated choice experiments for estimating nested logit models. *Transportation Research B*, Vol. 43, No. 1, pp. 19-35.

Bliemer, M.C.J., Rose, J.M., and Hess, S. (2008) Approximation of Bayesian efficiency in experimental choice designs. *Journal of Choice Modelling*, Vol. 1, No. 1, pp. 98-127.

Ben-Akiva, M., and Lerman, S. (1985) *Discrete choice analysis: theory and application to travel demand*. MIT Press, Cambridge, MA, USA.

Bunch, D.S., Louviere, J.J. and Anderson, D.A. (1994) A comparison of experimental design strategies for Multinomial Logit Models: the case of generic attributes. Working Paper, Graduate School of Management, University of California at Davis, CA, USA.

Burgess, L. and Street, D.J. (2003) Optimal designs for 2^k choice experiments. *Communications in Statistics, Theory and Methods*, Vol. 32, No. 11, pp. 2185–2206.

Carlsson, F. and Martinsson, P. (2003) Design techniques for stated preference methods in health economics. *Health Economics*, Vol. 12 (June), pp. 281-294.

Ferrini, S. and Scarpa, R. (2007) Designs with a-priori information for nonmarket valuation with choice-experiments: a Monte Carlo study. *Journal of Environmental Economics and Management*, Vol. 53, pp. 342-363.

Green, P.E., Krieger, A., and Agarwal, M.K. (1991) Adaptive conjoint analysis: Some caveats and suggestions. *Journal of Marketing Research*, Vol. 23, No., pp. 215–222.

Hauser, J.R., and Toubia, O. (2005) The impact of utility balance and endogeneity in conjoint analysis. *Marketing Science*, Vol. 24, No. 3, pp. 498-507.

Hensher, D.A., Rose, J.M., and Greene, W.H. (2005) *Applied choice analysis: a primer*. Cambridge University Press, Cambridge, UK.

Huber, J. and Zwerina, K. (1996) The importance of utility balance in efficient choice designs. *Journal of Marketing Research*, Vol. 33 (August), pp. 307-317.

Johnson, R. (1987) Accuracy of utility estimation in ACA. Working paper, Sawtooth Software, Sequim, WA, USA.

Johnson, R. (1991) Comment on "Adaptive conjoint analysis: Some caveats and suggestions." *Journal of Marketing Research*, Vol. 28(May), pp. 223-225.

Johnson, F.R., Kanninen, B.J. and Bingham, M. (2006) Experimental design for stated choice studies. In: Kanninen, B.J. (Ed.) *Valuing environmental amenities using stated choice studies: a common sense approach to theory and practice*. Springer, the Netherlands, pp. 159-202.

Judge, G.G., Griffiths, W.E., Hill, R.C., Lutkepohl, H., and Lee, T.C. (1985) *The theory and practice of econometrics*. John Wiley and Sons, New York, NY, USA.

Kanninen, B.J. (2002) Optimal design for multinomial choice experiments. *Journal of Marketing Research*, Vol. 39, pp. 214-217.

Kanninen, B.J. (2005) Optimal design for binary choice experiments with quadratic or interactive terms. Paper presented at the *International Health Economics Association Conference*, Barcelona, Spain.

Kessels, R., Goos, P. and Vandebroek, M. (2006) A comparison of criteria to design efficient choice experiments. *Journal of Marketing Research*, Vol. 43 (August), pp. 409-419.

Kessels, R., Jones, B., Goos, P., and Vandebroek, M. (in press), "An efficient algorithm for constructing Bayesian optimal choice designs. Forthcoming in *Journal of Business and Economic Statistics*.

Kuhfeld, W.F., Tobias, R.D., and Garratt, M. (1994) Efficient experimental design with marketing research applications. *Journal of Marketing Research*, Vol. 31, pp. 545-557.

Louviere, J.J., Hensher, D.A. and Swait, J.D. (2000), *Stated choice methods: analysis and application*. Cambridge University Press, Cambridge, UK.

McFadden, D. (1974) Conditional logit analysis of qualitative choice behaviour. In: Zarembka, P. (ed.) *Frontiers of econometrics*. Academic Press.

Rose, J.M., and Bliemer, M.C.J. (2008) Stated preference experimental design strategies. In: Hensher, D.A., and Button, K.J. (eds) *Handbook of Transport Modelling*. Elsevier, Oxford, UK, pp. 151-79.

Rose, J.M., Bliemer, M.C.J., Hensher, D.A., and Collins, A. (2008) Designing efficient stated choice experiments in the presence of reference alternatives. *Transportation Research Part B*, Vol. 42, pp. 395-406.

Sándor, Z. and Wedel, M. (2001) Designing conjoint choice experiments using managers' prior beliefs. *Journal of Marketing Research*, Vol. 38 (November), pp. 430-444.

Sándor, Z. and Wedel, M. (2002) Profile construction in experimental choice designs for mixed logit models. *Marketing Science*, Vol. 21, No. 4, pp. 455-475.

Sándor, Z. and Wedel, M. (2005) Heterogeneous conjoint choice designs. *Journal of Marketing Research*, Vol. 42 (May), pp. 210-218.

Street, D.J., Bunch, D.S. and Moore, B.J. (2001) Optimal designs for 2^k paired comparison experiments. *Communications in Statistics, Theory and Methods*, Vol. 30, No. 10, pp. 2149–2171.

Street, D.J. and Burgess, L. (2004) Optimal and near optimal pairs for the estimation of effects in 2-level choice experiments. *Journal of Statistical Planning and Inference*, Vol. 118, pp. 185–199.

Street, D., Burgess, L.B., and Louviere, J.J. (2005) Quick and easy choice sets: constructing optimal and nearly optimal stated choice experiments. *International Journal of Research in Marketing*, Vol. 22, No. 4, pp. 459-470.

Toner, J.P., Clark, S.D., Grant-Muller, S.M. and Fowkes, A.S. (1999) Anything you can do, we can do better: a provocative introduction to a new approach to stated preference design, *WCTR Proceedings*, Antwerp, Belgium, Vol. 3, pp. 107-120.

Toubia, O., Simester, D.L., Hauser, J.R., and Dahan, E. (2003) Fast polyhedral adaptive conjoint estimation. *Marketing Science*, Vol. 22, No. 3, pp. 273-303.

Train, K. (2005) *Discrete choice methods with simulation*. Cambridge University Press, Cambridge, UK.

Yu, J., Goos, P., and Vandebroek, M. (2009) Efficient conjoint choice designs in the presence of respondent heterogeneity. *Marketing Science*, Vol. 28, No. 1.