Optimal Designs for the Measurement of Consumer Preferences

Roselinde Kessels

Public defence - 23rd October 2006

Committee

Prof. Dr. M. Vandebroek (advisor)	*Katholieke Universiteit Leuven*
Prof. Dr. P. Goos (co-advisor)	*Universiteit Antwerpen*
Prof. Dr. C. Croux	*Katholieke Universiteit Leuven*
Prof. Dr. M. Dekimpe	*Universiteit van Tilburg &* *Katholieke Universiteit Leuven*
Dr. Z. Sándor	*Universidad Carlos III de Madrid*

Daar de proefschriften in de reeks van de Faculteit Economische en Toegepaste Economische Wetenschappen het persoonlijk werk zijn van hun auteurs, zijn alleen deze laatsten daarvoor verantwoordelijk.

New ideas come into this world somewhat like falling meteors, with a flash and an explosion, and perhaps somebody's castle-roof perforated.

Henry David Thoreau

Acknowledgements

The way to develop self-confidence is to do
the thing you fear and get a record of
successful experiences behind you.

William Jennings Bryan

I have always been eager to learn new things in all sorts of areas. When I obtained my undergraduate degree in Commercial Engineering at the K. U. Leuven, I felt like there was much more that I could learn at University. Particularly Lieven Tack, my Master's thesis instructor, gave me a flavor of life at University as a PhD student. When Martina Vandebroek, my Master's thesis advisor, asked me to start a PhD in design of marketing experiments, I simply could not refuse. A large desk in a bright office at the fifth floor with view over Leuven was reserved for me. Now, four years later, I really can testify that I have learned and experienced a lot. In all, doing research has shaped my mind in a more critical and creative way, which I think will prove useful in every aspect of life.

Although preparing a PhD resembles an adventure with considerable uncertainty at the start, many people have been vital to the fulfilment of this project. Foremost, I would like to express my sincere thanks to my advisor, Martina Vandebroek. Due to her profound insights, my first work led to a publication in the prestigious *Journal of Marketing Research*. This publication was a big boost to my confidence and encouraged me to continue my

PhD research with vigor. Martina has always been a careful reader of my work supporting me to write clearly and without any redundant complexity. Her comments and suggestions have contributed substantially to the quality of this dissertation. Also, she agreed on my visiting other universities and participating in conferences and workshops at many different places around the world. In that sense, I was perhaps not the easiest PhD student to deal with since I have constantly come up with new travelling ideas.

I am also very much indebted to my co-advisor, Peter Goos. Like Martina, he has always been there to support and guide me. At conferences people often talk about his excellent book and praise him for his outstanding writing talent and superb programming skills. At the same time, they are surprised that he is so young. I benefited greatly from his incredible expertise and enjoyed very much working with him. He also gave me the opportunity to visit the University of Antwerp where I joined his department for a month.

Furthermore, I am grateful to the other members of my doctoral committee: Christophe Croux, Marnik Dekimpe and Zsolt Sándor. Particularly, I have been delighted that Marnik Dekimpe accepted the very sudden invitation to become part of the committee. They all provided me with lots of provoking questions and insightful comments based on their refined and extended knowledge in their specific research domain.

In my last PhD year, I have been very privileged to get the chance to collaborate with Bradley Jones, an amazingly creative researcher at SAS/JMP. I owe him a great deal. He uncovered several mathematical principles for me and gave me a better understanding of how to use statistical tools. In addition, he has been a magnificent teacher to me in English writing and presentation skills. Before I met Bradley, I had taken several courses in English, but none can compete with the English I have learned from him. He also showed me around in the US and gave me the taste of Starbucks coffee, which I am drinking now every morning. If people were to ask me the name of a contemporary genius, then I would refer to Bradley. To give an idea about his tremendous expertise in design of experiments, I would like to quote Stuart Hunter, an experienced researcher himself, about Bradley in an

after dinner talk at a conference: *I will not talk about design of experiments because Bradley Jones is in the audience. Let me make myself clear through the following story. Once, there was a young guy who had survived a big flood and told everyone about it for the rest of his life. When he died and went to heaven he was about to tell God about it, but God stopped him and said, "Everyone gets a chance to tell his story. You will get yours too. But, just remember Noah is here."*

Next, I wish to thank Hans Nyquist to host me at his department of the University of Stockholm for six weeks. He also invited me for the yearly Winterconference of the Swedish statistical society held in Lapland. I had been very happy to leave Leuven for a while to get some fresh air in Sweden. And this has to be taken also literally because of the dust particles that were polluting the Belgian atmosphere when I left for Sweden early 2006. At the conference in Lapland, I was pleased to meet Gunnar Kulldorff immersing me in Swedish life and traditions and pointing me at some reindeer. After my Swedish adventure, he has constantly shown interest in the progress of my PhD work. Also, he has provided me with exciting updates on Sweden. I want to thank him for his original e-mails I have received which cheered me in my writing.

During the first years of my PhD research, I felt very fortunate to have such a great colleague as Arvind Ruggoo. He had always been there to help me around and was concerned about my well-being in general. We became good friends and had lots of fun during tea breaks, seminars and conferences. Due to Arvind's warm personality, we all soaked up a bit of the golden rays of the glorious Mauritian sun. I am very grateful to him and hope to visit him once in Mauritius.

Furthermore, I am much indebted to Yuri Goegebeur and Jan Adem for their professional advice and help on statistical issues and practical matters. I was able to enjoy Jan's unique organizing talent on a ski vacation. Also, I am much beholden to Linda Moonen, who has been a wonderful office mate for all these years. She provided me with lots of useful information and lent me her computer to run my programs during the night. As such, I am also

very grateful to the IT department of our faculty for putting some computer rooms at my disposal. Also, I have been very content to start the doctoral program with Aurélie Lemmens and Kristel Joossens. Aurélie has always assisted me with marketing related issues and Kristel with different kinds of software programs. Together with Linda Moonen we formed the alphabetic ORSTAT girls quartet "Joossens, Kessels, Lemmens and Moonen". In addition, I would like to embrace my other colleagues from the research groups ORSTAT and Operations Management for the congenial working atmosphere and the numerous evenings out in Leuven. Stefan's enthusiasm has always been catching, Roel's feel for languages has been admirable, and Olivier's advice on Stockholm and Jeroen's LaTeX style files came in very handy. Furthermore, a special word of thanks should go to my colleagues Bart Vermeulen and Jie Yu. Keep on the hard work, be patient and you will get there too!

Finally, I would like to thank all my friends for keeping my mind off work during weekends and evenings. Nevertheless, they have always been keen on listening to my adventures. I would like to mention my secondary school friends Ilse Beck and Evi Denolet, who are about to obtain a PhD degree as well. Good luck ladies! Above all, I am obliged to my parents who supported me in the most comprehensive way during all these years. Every weekend, they spoiled me with their 5-star services so that I rapidly got charged up again for another eventful week in Leuven.

Roselinde Kessels
Leuven, October 23, 2006

English preface

This thesis focuses on the design of conjoint experiments for measuring the trade-offs people make in choosing between alternative products and service providers. Marketing consultants and researchers frequently use these experiments to predict people's choices for prospective goods. In this way, they assist companies in launching innovative products or services. The entire process from collecting consumer preference data to analyzing them and simulating the marketplace is generally known as conjoint analysis.

Conjoint analysis assumes that a product or service can be decomposed into its component attributes and levels. For example, Table 1 contains three profiles or alternatives of a car that are described by levels for five attributes. By presenting a series of profiles to a number of test persons and finding out which are most preferred, conjoint analysis allows the determination of the relative importance of each attribute and level in the purchasing decision. The relative values or utilities respondents derive from the attribute levels are also called part-worths. Conjoint analysis is based on the fact that the part-worths can better be measured when the attributes are *con*sidered *joint*ly rather than in isolation. This is because the evaluation task is easier if the respondent is presented with combinations of attribute levels than with individual attributes and levels.

Respondents usually evaluate profiles in one of the following two ways. They either choose their preferred profile from a set of profiles, also called a choice set, and they repeat this task for several other choice sets presented to them. Or, they rate a number of profiles on a scale, for example a 10-point scale. The first type of conjoint experiment is a choice-based conjoint experiment, also referred to as a conjoint choice or discrete choice experiment, or more

Table 1: Three example profiles of a car.

Attribute	Profile 1	Profile 2	Profile 3
Price	18,000 EUR	17,000 EUR	14,000 EUR
Transmission	manual	automatic	manual
Airbags	front and dual side	front	front and dual side
Door lock	auto lock function	key-less entry remote control	key
Audio system	radio and CD player	radio and CD player and surround sound	radio

succinctly, a choice experiment. The second type of conjoint experiment is a rating-based conjoint experiment. These experiments have traditionally been used in conjoint analysis from its introduction in the seventies. Later on, in the eighties, conjoint choice experiments originated. They have become very popular because they are more realistic in imitating real shopping behavior which makes the task less difficult. However, they require more respondents since making choices is less informative than providing ratings.

In this thesis we deal with the question of how to properly design choice-based and rating-based conjoint experiments. This means that we search for profiles that, when administered to respondents, result in maximum information on the part-worths. To find the best possible design in each case, we make use of design criteria or optimality criteria resulting in optimal designs. The thesis is split into two parts each involving the design of one type of experiment. The first part consists of three chapters that concern the design of choice-based conjoint experiments. In Chapter 1 we discuss and compare four optimality criteria to construct choice designs. In Chapter 2 we develop an efficient algorithm to generate the designs. Chapter 3 concludes the part on choice designs providing practical recommendations on their use. The second part has two chapters focusing on the design of rating-based conjoint experiments. The first chapter, Chapter 4, describes an algorithmic construction of these designs, whereas the second chapter, Chapter 5, provides a manual building strategy addressing two-level designs for a large number of attributes. Each chapter starts with an introduction and ends with a conclusion in which avenues for future research are considered.

Dutch preface

Optimale ontwerpen voor het meten van consumentenvoorkeuren

Deze thesis handelt over het ontwerp van conjoint experimenten die gebruikt worden om inzicht te verwerven in de afwegingen van consumenten bij het kiezen van producten en diensten. Marketing consultants en onderzoekers voeren deze experimenten vaak uit om de voorkeuren voor toekomstige goederen te voorspellen. Op deze manier helpen ze bedrijven bij het lanceren van innovatieve producten of diensten. Het complete proces vanaf het verzamelen van gegevens met consumentenvoorkeuren tot het analyseren van de gegevens en het nabootsen van de markt staat bekend als conjoint analyse.

Conjoint analyse gaat uit van de veronderstelling dat een product of dienst kan ontbonden worden in een reeks van componenten of attributen die elk een bepaald niveau aannemen. Bijvoorbeeld, Tabel 2 toont drie profielen of alternatieven van een wagen die samengesteld zijn uit niveaus voor vijf attributen. Door een aantal proefpersonen een reeks profielen voor te leggen en hen te bevragen naar de meest aantrekkelijke profielen, kan met conjoint analyse het relatieve belang van elk attribuut en niveau in de aankoopbeslissing bepaald worden. De waarde die respondenten hechten aan elk van de attribuutniveaus wordt weergegeven door zogenaamde "part-worths". Conjoint analyse is gebaseerd op het feit dat de part-worths beter gemeten kunnen worden wanneer de attributen tezamen worden beschouwd (in het Engels wordt dit vertaald als "*con*sidered *joint*ly") dan elk afzonderlijk. Dit is omdat de evaluatietaak gemakkelijker is wanneer combinaties van attribuutniveaus worden voorgelegd in plaats van individuele attributen en niveaus.

Table 2: Drie voorbeeldprofielen van een wagen.

Attribuut	Profiel 1	Profiel 2	Profiel 3
Prijs	18.000 EUR	17.000 EUR	14.000 EUR
Transmissie	manueel	automatisch	manueel
Airbags	frontaal en lateraal	frontaal	frontaal en lateraal
Deurvergrendeling	automatisch	afstandsbediening zonder sleutel	sleutel
Audiosysteem	radio en CD-speler	radio en CD-speler met klankeffect	radio

Gewoonlijk evalueren respondenten profielen op één van de volgende twee manieren. Ofwel kiezen de respondenten het meest aantrekkelijke profiel uit een keuzeset van profielen, en doen ze dit voor verschillende keuzesets. Ofwel geven ze hun voorkeur voor alle profielen weer op een schaal, bijvoorbeeld een 10-puntenschaal. Het eerste type conjoint experiment is een keuze-gebaseerd conjoint experiment, ook wel een conjoint of discreet keuze-experiment genoemd, of simpelweg, een keuze-experiment. Het tweede type conjoint experiment is een score- of rating-gebaseerd conjoint experiment, of kortweg, een rating-experiment. Deze laatste experimenten werden traditioneel gebruikt in conjoint analyse sinds haar onstaan in de jaren '70. Later in de jaren '80 deden keuze-experimenten hun intrede. Keuze-experimenten zijn erg populair omdat ze een realistischer beeld geven van hoe het er in de markt aan toegaat. Dit maakt het voor de respondenten gemakkelijker om profielen te beoordelen. Daarentegen vereisen ze meer respondenten dan rating-experimenten omdat het maken van keuzes minder informatief is dan het verschaffen van scores.

In deze thesis houden we ons bezig met de vraag op welke manier keuze- en rating-experimenten het best ontworpen kunnen worden. Meer specifiek gaan we op zoek naar profielen die, wanneer ze voorgelegd worden aan proefpersonen, zoveel mogelijk informatie verschaffen over de part-worths. Om het best mogelijke ontwerp te vinden voor een bepaalde situatie maken we gebruik van ontwerpcriteria of optimaliteitscriteria. Hiermee kunnen optimale ontwerpen berekend worden. Deze thesis is opgesplitst in twee delen. Elk deel behandelt het ontwerp van een bepaald experiment. Het eerste deel bestaat uit

drie hoofdstukken die elk gaan over het ontwerp van keuze-experimenten. In Hoofdstuk 1 bespreken en vergelijken we vier optimaliteitscriteria die gebruikt kunnen worden voor de samenstelling van keuze-ontwerpen. In Hoofdstuk 2 ontwikkelen we een efficiënt algoritme om dergelijke ontwerpen te genereren. Hoofdstuk 3 sluit het deel over keuze-ontwerpen af met het verlenen van praktisch advies over hun gebruik. Het tweede deel omvat twee hoofdstukken die handelen over het ontwerp van rating-experimenten. Hoofdstuk 4 beschrijft een algoritmische opbouw van deze ontwerpen, terwijl Hoofdstuk 5 een manuele strategie voorstelt voor de samenstelling van ontwerpen die bestaan uit een groot aantal attributen met elk twee niveaus. Ieder hoofdstuk begint met een introductie en eindigt met een besluit waarin suggesties voor verder onderzoek geformuleerd worden.

Contents

Part I

Choice-based conjoint design

Chapter 1

A comparison of optimality criteria for choice designs

This chapter has been published as
↬ Kessels, R., Goos, P. and Vandebroek M. (2006). A comparison of criteria to design efficient choice experiments, *Journal of Marketing Research* **43**: 409–419.

A shorter Dutch version has also appeared as
↬ Kessels, R., Goos, P. and Vandebroek M. (2005). Het optimaal ontwerp van keuze-experimenten, *Business InZicht* **20**: 2–3.

Abstract

To date, no attempt has been made to design efficient conjoint choice experiments by means of the $\mathcal{G}$- and $\mathcal{V}$-optimality criteria. These criteria are known to make precise response predictions which is exactly what conjoint choice experiments aim to do. In this chapter, we elaborate on the $\mathcal{G}$- and $\mathcal{V}$-optimality criteria for the multinomial logit model and compare their prediction performances with those of the $\mathcal{D}$- and $\mathcal{A}$-optimality criteria. We make use of Bayesian design methods that integrate the optimality criteria over a prior distribution of likely parameter values. A modified Fedorov algorithm is employed to generate the optimal choice designs. Other aspects of the designs, such as level overlap, utility balance, estimation performance and computational effectiveness, are also discussed.

1.1 Introduction

Since Louviere and Woodworth's (1983) article, conjoint choice experiments or more succinctly, choice experiments, have become increasingly popular to explore consumer preferences for the attributes of various goods. In applied research, these experiments have been used extensively, and in fundamental research, they have been the subject of rigorous study and research. The reason for their popularity is that they enable researchers to model real marketplace choices and thus to emulate real market decisions and predict market demand (Carson et al. 1994). In a typical choice experiment, respondents are presented with a series of choice sets, each composed of several alternatives, also called profiles, of products or services that are defined as combinations of different attribute levels. Respondents then indicate their preferred alternative for every choice set.

Louviere, Street and Burgess (2003) present an overview of the recent developments in choice experiments, with a special emphasis on the design of these experiments. Designing an efficient choice experiment involves selecting alternatives that, when put into choice sets, provide maximum information on the parameters of a probabilistic choice model. Until now, the efficiency of a choice design has been expressed primarily in terms of the $\mathcal{D}$-optimality criterion (Atkinson and Donev 1992). Only Street, Bunch and Moore (2001) applied the $\mathcal{A}$-optimality criterion to the design of paired comparison experiments with two-level attributes. In a paired comparison design, every choice set consists of two alternatives. To date, the $\mathcal{G}$- and $\mathcal{V}$-optimality criteria, specifically developed for making precise response predictions, have not been applied in the experimental choice context. However, choice experiments are conducted for predictive purposes, and therefore, we turn attention to the $\mathcal{G}$- and $\mathcal{V}$-optimality criteria.

The main difficulty in the construction of a proper choice design is that the probabilistic choice models are nonlinear in the parameters, implying that the efficiency of the design depends on the unknown parameter vector (Atkinson and Haines 1996). Consequently, researchers need to assume values for the parameters before deriving the experimental design. To circumvent this circular problem, three approaches have been introduced. We discuss them for

logit choice models, the best known of which is the multinomial logit model (McFadden 1974).

The first approach is to use zero prior parameter values so that methods of linear experimental design can be applied. It is implicitly assumed that the respondents prefer all attribute levels and, thus, all alternatives equally (Grossmann, Holling and Schwabe 2002). The following authors are representatives of this approach. Anderson and Wiley (1992) and Lazari and Anderson (1994) provided a catalog of orthogonal arrays for logit choice models. To address a broader range of design classes, Kuhfeld, Tobias and Garratt (1994) made use of Cook and Nachtsheim's (1980) modification of Fedorov's (1972) exchange algorithm to generate $\mathcal{D}$-optimal designs. Kuhfeld and Tobias (2005) continued this line of research by integrating the modified Fedorov algorithm in a comprehensive algorithm, contained in the SAS `%MktEx` macro, which also exploits a large catalog of orthogonal arrays and Meyer and Nachtsheim's (1995) coordinate-exchange algorithm.

Furthermore, Bunch, Louviere and Anderson (1996) developed the so-called $\mathcal{D}$-optimal shifted or cyclic designs characterized by the minimal level overlap property. This property is satisfied when the frequencies of the attribute levels within a choice set are distributed as equally as possible. The results for $\mathcal{D}$-optimal paired comparison designs for two-level attributes have been described in the work of Street, Bunch and Moore (2001), Street and Burgess (2004) and references therein. As we mentioned previously, Street, Bunch and Moore (2001) also computed $\mathcal{A}$-optimal paired comparison designs for two-level attributes. Finally, Burgess and Street (2003) derived $\mathcal{D}$-optimal choice designs for two-level attributes of any choice set size and extended these results in Burgess and Street (2005) to apply to attributes with any number of levels.

The second approach, attributed to the work of Huber and Zwerina (1996), advocates the use of nonzero prior values rather than zero values. The resulting locally $\mathcal{D}_P$-optimal designs prove to be more efficient than the $\mathcal{D}$-optimal designs based on zero prior values. Carlsson and Martinsson (2003) confirmed this finding with a comparison study in health economics. To generate the

$\mathcal{D}_P$-optimal designs, Huber and Zwerina (1996) proposed the relabeling (R) and swapping (S) techniques, shortly referred to as the RS algorithm. The SAS `%ChoicEff` macro that uses a modified Fedorov algorithm also allows building the designs, as illustrated by Zwerina, Huber and Kuhfeld (1996; see updated [2005] version).

Finally, the most recent approach has been introduced by Sándor and Wedel (2001) and consists of integrating the associated uncertainty on the assumed parameter values by the use of Bayesian design techniques (Chaloner and Verdinelli 1995). If there is substantial uncertainty about the unknown parameters, the so-called Bayesian $\mathcal{D}_B$-optimal designs outperform the locally $\mathcal{D}_P$-optimal designs. The algorithm used is the RS algorithm and an additional cycling (C) procedure, accordingly called the RSC algorithm. Sándor and Wedel (2002) developed an updated version of this algorithm.

The foregoing researchers have proposed designs for the multinomial logit model to be administered to various respondents whose choices are pooled. As a result, homogeneous parameters across respondents are assumed. In this chapter, we adopt the same experimental choice scenario to compare the performances of the $\mathcal{D}$-, $\mathcal{A}$-, $\mathcal{G}$- and $\mathcal{V}$-optimality criteria. Note that we study main-effects choice designs only. Our approach is similar to that of Sándor and Wedel (2001) in that we also implement Bayesian design methods. However, we do not apply the RSC algorithm but rather the modified Fedorov algorithm to generate the optimal designs.

The outline of the remainder of this chapter is as follows. Section 1.2 describes the multinomial logit model and the $\mathcal{D}$-, $\mathcal{A}$-, $\mathcal{G}$- and $\mathcal{V}$-optimality criteria using this model. The approach to generate the optimal designs with the modified Fedorov algorithm is discussed in Section 1.3. In Section 1.4, the different optimal designs are presented and compared. Finally, Section 1.5 concludes the chapter and provides suggestions for future research.

1.2 Multinomial logit and optimality criteria

The multinomial logit model is derived from McFadden's (1974) random utility model of consumer choice. In the random utility model, the utility a person attaches to a given profile j is predicted by

$$U_j = \mathbf{x}_j'\boldsymbol{\beta} + \varepsilon_j, \tag{1.1}$$

where $\mathbf{x}_j$ is a $k \times 1$ vector of the attribute levels of profile j, $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_k]'$ is a $k \times 1$ vector of parameter values or part-worths weighing the attribute levels and ε_j is an i.i.d. extreme value error term.

Now, consider presenting N respondents with a choice experiment containing S choice sets, $s = 1, 2, ..., S$, where each choice set consists of J profiles, $j = 1, 2, ..., J$. Each respondent indicates the profile that maximizes that respondent's perceived utility for every choice set. The multinomial logit probability that a respondent chooses profile j in choice set s is

$$p_{js}(\mathbf{X}_s, \boldsymbol{\beta}) = \frac{e^{\mathbf{x}_{js}'\boldsymbol{\beta}}}{\sum_{t=1}^{J} e^{\mathbf{x}_{ts}'\boldsymbol{\beta}}}, \tag{1.2}$$

where $\mathbf{X}_s = [\mathbf{x}_{1s}, ..., \mathbf{x}_{Js}]'$ is the design matrix for choice set s. The stacked $\mathbf{X}_s$ matrices provide the design matrix $\mathbf{X}$ for the choice experiment.

Because of the assumption of independent error terms, the choices from the N respondents in the S choice sets represent independent draws from a multinomial distribution. Therefore, if $\mathbf{Y} = [\mathbf{y}_1, ..., \mathbf{y}_N]$ denotes the matrix of choices from the N respondents with elements y_{jsn}, each of which equals one if respondent n, $n = 1, 2, ..., N$, chooses alternative j in choice set s and zero otherwise, then the log-likelihood of the N samples in $\mathbf{Y}$ is defined by

$$LL(\mathbf{Y}|\mathbf{X}, \boldsymbol{\beta}) = \sum_{s=1}^{S}\sum_{j=1}^{J}\sum_{n=1}^{N} y_{jsn} \ln\left(p_{js}(\mathbf{X}_s, \boldsymbol{\beta})\right). \tag{1.3}$$

Maximizing this expression with respect to $\boldsymbol{\beta}$ yields the maximum likelihood estimator $\hat{\boldsymbol{\beta}}$ for a particular choice design.

The construction of efficient choice designs is based on the Fisher information matrix, which is the inverse of the variance-covariance matrix of the parameter estimators, and given by

$$\mathbf{M}(\mathbf{X}, \boldsymbol{\beta}) = N \sum_{s=1}^{S} \mathbf{X}_s'(\mathbf{P}_s - \mathbf{p}_s\mathbf{p}_s')\mathbf{X}_s, \tag{1.4}$$

where $\mathbf{p}_s = [p_{1s}, ..., p_{Js}]'$ and $\mathbf{P}_s = \text{diag}[p_{1s}, ..., p_{Js}]$. In Appendix A, we show how the information matrix is obtained from the log-likelihood function (1.3). In optimal design theory (Atkinson and Donev 1992; Fedorov 1972; Silvey 1980), direct functions of the information matrix, referred to as optimality criteria or design criteria, are proposed to generate optimal designs that yield precise parameter estimates or accurate predictions. However, the information matrix (1.4) depends on the unknown parameters through the probabilities so that parameter values are required before constructing optimal choice designs. As we mentioned in Section 1.1, Sándor and Wedel (2001) adopted a Bayesian design approach that involves the specification of a prior parameter distribution $\pi(\boldsymbol{\beta})$. Usually, this distribution is the normal distribution, $\mathcal{N}(\boldsymbol{\beta}|\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$, from which R prior parameter values $\boldsymbol{\beta}^r$, $r = 1, ..., R$, are drawn to approximate it. In general, the resulting Bayesian optimal designs outperform locally optimal designs that are based on a single prior parameter.

1.2.1 $\mathcal{D}$- and $\mathcal{A}$-optimality criteria

The most popular optimality criterion to design choice experiments is the $\mathcal{D}$-optimality criterion. The $\mathcal{D}$-optimality criterion seeks to maximize the determinant of the information matrix (1.4), or to minimize its inverse, the determinant of the variance-covariance matrix of the parameter estimators. It is related to the $\mathcal{A}$-optimality criterion that prefers the design for which the sum or the average of the variances of the parameter estimators is minimized. However, a drawback of the $\mathcal{A}$-optimality criterion is that the ordering of designs with respect to this criterion depends on the type of coding. We refer to the work of Goos (2002, pages 38–40) for an example in the case of linear models. Note that the $\mathcal{A}$-optimality criterion is more suited for obtaining precise parameter estimates because it considers the variances of the estimators only. In addition, the $\mathcal{D}$-optimality criterion takes the covariances into

account. Formally, the Bayesian $\mathcal{D}$-criterion value, the $\mathcal{D}_B$-criterion value, is

$$\mathcal{D}_B = \int_{\mathcal{R}^k} \left\{\det(\mathbf{M}^{-1}(\mathbf{X}, \boldsymbol{\beta}))\right\}^{1/k} \pi(\boldsymbol{\beta}) \mathrm{d}\boldsymbol{\beta}, \tag{1.5}$$

where the exponent $1/k$ ensures that it is independent of the dimension k of the parameter vector $\boldsymbol{\beta}$. Minimizing this value results in the $\mathcal{D}_B$-optimal design. The $\mathcal{A}_B$-optimal design minimizes

$$\mathcal{A}_B = \int_{\mathcal{R}^k} \mathrm{tr}(\mathbf{M}^{-1}(\mathbf{X}, \boldsymbol{\beta}))\pi(\boldsymbol{\beta}) \mathrm{d}\boldsymbol{\beta}. \tag{1.6}$$

1.2.2 $\mathcal{G}$- and $\mathcal{V}$-optimality criteria

The $\mathcal{G}$- and $\mathcal{V}$-optimality criteria both look for designs that make precise predictions about the response. Although choice experiments are carried out to make precise predictions about consumers' future purchasing behavior, the $\mathcal{G}$- and $\mathcal{V}$-optimality criteria have not yet been applied in the experimental choice context. To derive the $\mathcal{G}$- and $\mathcal{V}$-optimality criteria for the nonlinear choice model, predicted probabilities must be computed, and to do so, choice sets must be specified. In particular, we computed the predicted probabilities with respect to all possible choice sets of size J that can be composed from the candidate profiles. These choice sets make up the so-called design region χ. Thus, if there are Q possible choice sets, $\chi = \{\{\mathbf{x}_{1q}, ..., \mathbf{x}_{Jq}\} | \; q = 1, ..., Q\}$. Then, by definition, the $\mathcal{G}$-optimality criterion aims to minimize the maximum prediction variance over the design region χ, whereas the $\mathcal{V}$-optimality criterion aims to minimize the average prediction variance over this region. Mathematically, the $\mathcal{G}_B$-criterion value is given by

$$\begin{aligned} \mathcal{G}_B &= \int_{\mathcal{R}^k} \max_{\mathbf{x}_{jq} \in \chi} \mathrm{var}\{\hat{p}_{jq}(\mathbf{x}_{jq}, \boldsymbol{\beta})\}\pi(\boldsymbol{\beta}) \mathrm{d}\boldsymbol{\beta}, \\ &= \int_{\mathcal{R}^k} \max_{\mathbf{x}_{jq} \in \chi} \mathbf{c}'(\mathbf{x}_{jq}) \mathbf{M}^{-1}(\mathbf{X}, \boldsymbol{\beta}) \mathbf{c}(\mathbf{x}_{jq}) \pi(\boldsymbol{\beta}) \mathrm{d}\boldsymbol{\beta}, \end{aligned} \tag{1.7}$$

where $\hat{p}_{jq}(\mathbf{x}_{jq}, \boldsymbol{\beta})$ denotes the predicted choice probability for $\mathbf{x}_{jq}$ and

$$\mathbf{c}(\mathbf{x}_{jq}) = \frac{\partial p_{jq}(\mathbf{x}_{jq}, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}}, \tag{1.8}$$

the first-order truncated Taylor series expansion of the multinomial logit probability (1.2). This approach is similar to the computation of locally $\mathcal{D}$-

and c-optimal designs for nonlinear models in general (Atkinson and Donev 1992; Atkinson and Haines 1996). Using the multinomial logit model (1.2), we can write (1.8) as

$$\begin{aligned}
\mathbf{c}(\mathbf{x}_{jq}) &= \frac{e^{\mathbf{x}'_{jq}\boldsymbol{\beta}}\mathbf{x}_{jq}\sum_{t=1}^{J}e^{\mathbf{x}'_{tq}\boldsymbol{\beta}} - e^{\mathbf{x}'_{jq}\boldsymbol{\beta}}\sum_{t=1}^{J}e^{\mathbf{x}'_{tq}\boldsymbol{\beta}}\mathbf{x}_{tq}}{\left(\sum_{t=1}^{J}e^{\mathbf{x}'_{tq}\boldsymbol{\beta}}\right)^2}, \\
&= \frac{e^{\mathbf{x}'_{jq}\boldsymbol{\beta}}}{\sum_{t=1}^{J}e^{\mathbf{x}'_{tq}\boldsymbol{\beta}}}\left(\frac{\mathbf{x}_{jq}\sum_{t=1}^{J}e^{\mathbf{x}'_{tq}\boldsymbol{\beta}} - \sum_{t=1}^{J}e^{\mathbf{x}'_{tq}\boldsymbol{\beta}}\mathbf{x}_{tq}}{\sum_{t=1}^{J}e^{\mathbf{x}'_{tq}\boldsymbol{\beta}}}\right), \\
&= p_{jq}\left(\mathbf{x}_{jq} - \sum_{t=1}^{J}\left(\frac{e^{\mathbf{x}'_{tq}\boldsymbol{\beta}}}{\sum_{v=1}^{J}e^{\mathbf{x}'_{vq}\boldsymbol{\beta}}}\right)\mathbf{x}_{tq}\right), \\
&= p_{jq}\left(\mathbf{x}_{jq} - \sum_{t=1}^{J}p_{tq}\mathbf{x}_{tq}\right).
\end{aligned} \tag{1.9}$$

Akin to the $\mathcal{A}$-optimality criterion, the relative design efficiency in terms of the $\mathcal{G}$-optimality criterion is contingent on the type of coding. The same applies to the $\mathcal{V}$-optimality criterion, the Bayesian value of which is obtained as

$$\mathcal{V}_B = \int_{\mathcal{R}^k}\int_{\chi}\mathbf{c}'(\mathbf{x}_{jq})\mathbf{M}^{-1}(\mathbf{X},\boldsymbol{\beta})\mathbf{c}(\mathbf{x}_{jq})\mathrm{d}\mathbf{x}_{jq}\pi(\boldsymbol{\beta})\mathrm{d}\boldsymbol{\beta}, \tag{1.10}$$

with $\mathbf{c}(\mathbf{x}_{jq})$ defined in (1.8) and (1.9).

In Appendix B, we present a simple numerical example of the construction of Bayesian optimal designs by means of the $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimality criteria.

1.3 Algorithmic approach

The most embedded algorithms in the literature to generate choice designs are the RSC algorithm, embracing relabeling, swapping and cycling, and the modified Fedorov algorithm. There are two versions of the RSC algorithm developed by Sándor and Wedel, one in 2001 and one in 2002. As opposed to the first version, the updated version does not restrict its searches to designs

that satisfy the minimal level overlap property, provided that the starting design complies with it. This makes the RSC algorithm more prone to design improvements. As a result to its modification, the RSC algorithm generates designs that are statistically as efficient as those produced by the modified Fedorov algorithm. In the modified Fedorov algorithm, design profiles are exchanged with the profiles from a predefined set of candidate profiles without the enforcement of any constraint. We prefer the modified Fedorov algorithm because it is faster than the adjusted RSC algorithm in generating Bayesian optimal designs. Thus, we incorporated the $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimality criteria in the modified Fedorov algorithm to obtain four distinct Bayesian modified Fedorov choice algorithms. To avoid poor local optima, we repeated each of the algorithms for several starting designs. We refer to each repetition as a try and we performed 200 tries.

With the Bayesian modified Fedorov choice algorithms, we constructed $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs of two classes. The first class is given by designs of type $3^2 \times 2/2/12$— that is, designs with 12 choice sets, each of size two, in which each alternative is described by three attributes. The first two attributes have three levels each and the third attribute has two levels. The designs in the second class are of type $3^2 \times 2/3/8$, comprising 8 choice sets of size three and a similar attribute structure as the first design class. As a result, the sets of candidate profiles of both design classes are identical, enclosing the same $3^2 \times 2 = 18$ profiles. In addition, the designs of the two classes consist of the same number of profiles (i.e., 24) to compare the two- and three-alternative optimal designs with respect to specific design measures (see Section 1.4). To compute the $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs, the design region χ needs to be specified for each class. For the two-alternative design class, χ consists of $Q = \binom{18}{2} = 153$ choice sets, or 306 profiles, whereas for the three-alternative design class, it includes $Q = \binom{18}{3} = 816$ choice sets, or 2,448 profiles.

Furthermore, through the use of effects-type coding (see Appendix B), the number of parameter values, k, is five. As prior parameter distribution, we used the multivariate normal distribution $\pi(\boldsymbol{\beta}) = \mathcal{N}(\boldsymbol{\beta}|\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$, with $\boldsymbol{\beta}_0 = [-1, 0, -1, 0, -1]'$ and $\boldsymbol{\Sigma}_0 = \mathbf{I}_5$. The $\boldsymbol{\beta}_0$ vector is special because the

values for the levels of each of the attributes are equally spaced between -1 and 1. Through this scaling, the utilities increase with the levels of each attribute. For example, for the first two attributes that possess three levels each, a utility of -1 is attached to level 1, a utility of 0 to level 2 and a utility of 1 to level 3. A more extensive account on the specification of $\boldsymbol{\beta}_0$ can be found in the work of Huber and Zwerina (1996). Following Sándor and Wedel's (2001) example, we drew $R = 1,000$ samples $\boldsymbol{\beta}^r$ from $\pi(\boldsymbol{\beta})$.

1.4 Results

In this section, we compare the two- and three-alternative $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs with respect to their performances on several measures whilst taking into account some computational aspects. We performed all computations with the SAS 8.02 procedure IML (Interactive Matrix Language).

1.4.1 Performance of the $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimality criteria

We begin by illustrating the two- and three-alternative $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs, followed by a study of their amount of level overlap and degree of utility balance. We then score the robustness of the designs on other design criteria for which they are not optimized. Finally, we discuss the accuracy and predictive validity of the parameter estimates.

Designs, overlap and utility balance

The two- and three-alternative Bayesian optimal designs appear in Tables 1.1 and 1.2, respectively. Their criterion values appear in Table 1.3. The designs clearly exhibit some level overlap. As in the work of Sándor and Wedel (2002), we computed the percentage of the cases in which the columns of the choice sets do not satisfy the minimal level overlap property. The results appear in Table 1.4. The $\mathcal{D}_B$-optimal designs have the lowest level overlap in the two design classes, followed by the $\mathcal{V}_B$-optimal designs. In contrast, the

$\mathcal{G}_B$- and $\mathcal{A}_B$-optimal designs have the highest level overlap.

Table 1.1: Two-alternative Bayesian optimal designs.

		$\mathcal{D}_B$			$\mathcal{A}_B$			$\mathcal{G}_B$			$\mathcal{V}_B$		
Choice	Alt	Attr			Attr			Attr			Attr		
set		1	2	3	1	2	3	1	2	3	1	2	3
1	I	2	3	1	3	2	1	3	1	2	2	2	2
	II	1	2	1	3	1	1	2	2	2	1	1	1
2	I	2	2	2	2	3	1	3	2	1	2	1	2
	II	1	1	1	1	2	1	2	3	1	1	2	1
3	I	1	2	2	2	1	2	1	2	1	1	2	2
	II	3	1	2	1	2	2	2	1	1	3	1	1
4	I	2	2	1	3	1	1	1	3	1	2	2	1
	II	1	3	1	2	2	1	3	1	2	1	3	2
5	I	2	2	1	2	2	1	3	3	2	2	1	1
	II	3	3	2	1	1	1	2	1	1	3	2	2
6	I	2	1	2	3	3	2	2	3	1	1	3	1
	II	1	2	1	2	1	1	1	1	1	2	1	1
7	I	1	1	2	1	3	1	1	3	1	1	2	1
	II	2	2	2	2	2	1	2	1	1	3	3	2
8	I	1	2	2	1	3	1	2	2	2	3	2	1
	II	2	1	1	3	1	2	1	1	1	2	3	1
9	I	3	2	1	3	2	1	3	2	2	3	2	2
	II	2	1	1	2	3	2	1	2	2	1	1	2
10	I	3	1	1	1	2	2	1	2	1	2	3	1
	II	2	3	2	2	1	1	1	3	2	1	1	1
11	I	1	3	1	1	3	2	3	2	1	2	3	1
	II	3	1	2	3	3	2	2	2	2	2	2	2
12	I	2	1	1	2	1	2	1	2	2	3	1	2
	II	3	2	2	1	1	1	2	2	1	2	2	2

Table 1.2: Three-alternative Bayesian optimal designs.

		$\mathcal{D}_B$			$\mathcal{A}_B$			$\mathcal{G}_B$			$\mathcal{V}_B$		
Choice	Alt	Attr			Attr			Attr			Attr		
set		1	2	3	1	2	3	1	2	3	1	2	3
1	I	3	2	1	1	3	2	2	2	1	1	2	2
	II	2	1	1	2	3	1	3	1	2	3	1	1
	III	1	2	2	1	2	1	1	3	1	2	2	1
2	I	1	1	1	1	1	1	3	2	2	3	2	1
	II	2	2	1	2	1	1	1	1	2	1	3	1
	III	1	3	2	1	2	1	3	1	2	2	1	1
3	I	1	3	1	3	1	1	2	3	2	2	1	1
	II	2	3	2	2	2	1	1	2	1	1	2	1
	III	2	1	1	2	2	2	2	3	1	1	1	2
4	I	3	1	2	1	1	1	3	3	2	3	3	2
	II	1	2	1	1	3	2	3	1	1	2	2	2
	III	2	1	1	2	2	2	2	2	2	2	1	1
5	I	3	1	1	1	3	2	2	1	1	1	2	1
	II	3	2	2	2	1	2	2	3	2	2	3	1
	III	1	2	1	3	2	2	3	2	2	3	3	2
6	I	2	1	2	2	2	1	1	3	1	2	3	2
	II	1	2	1	3	2	2	2	3	2	3	1	2
	III	2	3	1	1	1	1	2	1	2	1	3	1
7	I	3	2	2	1	2	1	1	2	2	3	3	1
	II	1	1	1	2	1	2	2	1	2	3	2	2
	III	2	3	2	2	2	2	3	1	2	1	3	1
8	I	1	3	1	3	1	1	1	2	2	2	3	2
	II	2	2	1	1	3	1	1	1	1	1	1	1
	III	3	1	1	3	3	1	2	2	1	2	2	1

Table 1.3: $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-criterion values of the two- and three-alternative $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs.

Optimal design	# Alternatives	
	2	3
$\mathcal{D}_B$	0.73024	0.76617
$\mathcal{A}_B$	6.60563	6.02261
$\mathcal{G}_B$	0.51997	0.51843
$\mathcal{V}_B$	0.07219	0.06285

Table 1.4: Percentages of level overlap in the two- and three-alternative Bayesian optimal designs.

Optimal design	# Alternatives	
	2	3
$\mathcal{D}_B$	14%	38%
$\mathcal{A}_B$	28%	63%
$\mathcal{G}_B$	28%	54%
$\mathcal{V}_B$	17%	42%

To measure the utility balance of the computed designs, we built on the cumulative entropy of a choice design, as suggested by Swait and Adamowicz (2001). Utility balance is a concept that Huber and Zwerina (1996) introduced and it refers to the situation in which respondents prefer the alternatives in a choice set equally and thus face a difficult choice task. In the Bayesian framework, the cumulative entropy of a choice design is defined as

$$CH(\mathbf{X}, \boldsymbol{\beta}) = -\sum_{s=1}^{S} \int_{\mathcal{R}^k} \left(\sum_{j=1}^{J} \left\{ p_{js}(\mathbf{X}_s, \boldsymbol{\beta}) \ln(p_{js}(\mathbf{X}_s, \boldsymbol{\beta})) \right\} \right) \pi(\boldsymbol{\beta}) \mathrm{d}\boldsymbol{\beta}. \quad (1.11)$$

To derive the lower and upper bounds for the cumulative entropy for the two design classes, we constructed for each class a Bayesian design that is not at all utility balanced, or minimum utility balanced, and a Bayesian design that is maximum utility balanced. We obtained the former design by selecting the S choice sets that produced the smallest Bayesian cumulative entropy out

of all possible ones, whereas we initially constructed the latter by choosing the S choice sets that produced the largest Bayesian cumulative entropy. This could easily be done by enumerating all Q possible choice sets of size J. However, in doing so, the Bayesian maximum utility balanced designs turned out to be singular. We solved this problem by optimally replacing a minimum number of choice sets with choice sets with a slightly smaller Bayesian entropy.

The Bayesian minimum and maximum utility-balanced designs for the two design classes appear in Tables C1 and C2 of Appendix C together with their efficiencies with respect to the different optimality criteria. The minimum utility-balanced designs exhibit only a small amount of level overlap, whereas the maximum utility-balanced designs are characterized by a great deal of level overlap. The efficiencies of the designs are very low, particularly those of the minimum utility-balanced designs which are essentially zero. For the two-alternative design class, the values for the cumulative entropy of the Bayesian minimum and maximum utility-balanced designs are equal to 1.93 and 5.53, respectively. So these are the minimum and maximum values for the Bayesian cumulative entropy in the two-alternative design case. For the three-alternative design class, these values amount to 2.00 and 5.97.

Subsequently, we computed the values of the cumulative entropy for the two- and three-alternative Bayesian optimal designs and compared them with their maximum value. The values and their percentages appear in Table 1.5. On the whole, the designs are not maximum utility balanced but entail a moderate choice task complexity. This finding is counter to Huber and Zwerina's (1996) statement that proper choice designs must be maximum utility balanced. Although not perfectly utility balanced, the $\mathcal{A}_B$-optimal designs display the largest cumulative entropy, which extends Arora and Huber's (2001) result that $\mathcal{A}_P$-optimal designs for binary logit models are utility balanced to a Bayesian context. Conversely, the $\mathcal{V}_B$-optimal designs, which are developed especially for making precise predictions, exhibit the smallest cumulative entropy or the least complicated choice tasks. In addition, the three-alternative designs appear to be more complex than the two-alternative ones.

Table 1.5: Values and percentage values of cumulative entropy of the two- and three-alternative Bayesian optimal designs.

Optimal design	# Alternatives			
	2		3	
$\mathcal{D}_B$	3.98	72%	4.41	74%
$\mathcal{A}_B$	4.27	77%	4.70	79%
$\mathcal{G}_B$	3.96	72%	4.55	76%
$\mathcal{V}_B$	3.72	67%	4.29	72%

Performance in terms of other optimality criteria

Because the $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimality criteria all have a different aim, it is interesting to observe how robust the Bayesian optimal designs are to other design criteria for which they are not optimized. Panels a and b of Table 1.6 give the efficiencies of the two- and three-alternative optimal designs with respect to the different optimality criteria. As we expected from optimal design theory, the efficiencies of the $\mathcal{D}_B$-optimal designs on the $\mathcal{A}_B$-optimality criterion and of the $\mathcal{A}_B$-optimal designs on the $\mathcal{D}_B$-optimality criterion are quite high. This interdependence of criterion efficiencies also occurs between the $\mathcal{G}_B$- and $\mathcal{V}_B$-optimality criteria. Furthermore, compared with the $\mathcal{D}_B$-optimal designs, the $\mathcal{A}_B$-optimal designs do not score well in terms of $\mathcal{G}_B$- and $\mathcal{V}_B$-efficiency. As a result, the predictive ability of the $\mathcal{A}_B$-optimal designs is relatively low.

Accuracy and predictive validity of the parameter estimates

We now examine more closely the accuracy and predictive validity of the parameter estimates produced by the two- and three-alternative Bayesian optimal designs. To this end, we investigate the expected mean squared errors of the parameter estimates, $EMSE_{\hat{\boldsymbol{\beta}}}$, and of the predicted probabilities, $EMSE_{\hat{\mathbf{p}}_c}$. Both measures depend on a true parameter $\boldsymbol{\beta}_t$. The $EMSE_{\hat{\boldsymbol{\beta}}}$ pertains to the accuracy of the parameter estimates and is given by

$$EMSE_{\hat{\boldsymbol{\beta}}}(\boldsymbol{\beta}_t) = \int_{\mathcal{R}^k} \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_t\right)' \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_t\right) f(\hat{\boldsymbol{\beta}}) \mathrm{d}\hat{\boldsymbol{\beta}}, \tag{1.12}$$

Table 1.6: Performances of the Bayesian optimal designs in terms of other design criteria.

Evaluation criterion	a) Two-alternative designs			
	$\mathcal{D}_B$	$\mathcal{A}_B$	$\mathcal{G}_B$	$\mathcal{V}_B$
$\mathcal{D}_B$	100.00%	90.82%	89.08%	93.28%
$\mathcal{A}_B$	97.59%	100.00%	92.43%	87.13%
$\mathcal{G}_B$	94.49%	85.68%	100.00%	99.68%
$\mathcal{V}_B$	96.95%	88.12%	96.03%	100.00%

Evaluation criterion	b) Three-alternative designs			
	$\mathcal{D}_B$	$\mathcal{A}_B$	$\mathcal{G}_B$	$\mathcal{V}_B$
$\mathcal{D}_B$	100.00%	93.49%	94.13%	96.36%
$\mathcal{A}_B$	94.03%	100.00%	96.63%	89.80%
$\mathcal{G}_B$	80.81%	80.19%	100.00%	95.04%
$\mathcal{V}_B$	95.65%	86.04%	96.24%	100.00%

where $f(\hat{\boldsymbol{\beta}})$ is the distribution of the estimates. The smaller the $EMSE_{\hat{\boldsymbol{\beta}}}$ value, the more accurately the parameters are estimated. The $EMSE_{\hat{\mathbf{p}}_c}$ concerns the predictions with respect to the design that contains all Q possible choice sets of size J. This design is chosen so as not to favor any optimal design and is referred to as the complete choice design. It contains the same Q choice sets as the design region χ we defined in Sections 1.2.2 and 1.3. To compare the prediction performances of the two- and three-alternative optimal designs, we averaged the $EMSE_{\hat{\mathbf{p}}_c}$ values over the number of profiles in the complete choice design. Formally,

$$EMSE_{\hat{\mathbf{p}}_c}(\boldsymbol{\beta}_t) = \frac{1}{J \times Q} \int_{\mathcal{R}^k} \left(\hat{\mathbf{p}}_c(\hat{\boldsymbol{\beta}}) - \mathbf{p}_c(\boldsymbol{\beta}_t)\right)' \left(\hat{\mathbf{p}}_c(\hat{\boldsymbol{\beta}}) - \mathbf{p}_c(\boldsymbol{\beta}_t)\right) f(\hat{\boldsymbol{\beta}}) \mathrm{d}\hat{\boldsymbol{\beta}}, \tag{1.13}$$

where $\mathbf{p}_c(\boldsymbol{\beta}_t)$ is the vector of true logit probabilities in the complete choice design and $\hat{\mathbf{p}}_c(\hat{\boldsymbol{\beta}})$ is the corresponding vector of predicted logit probabilities. The smaller the $EMSE_{\hat{\mathbf{p}}_c}$ value, the more precisely the probabilities are predicted.

To approximate the distribution of the parameter estimates, $f(\hat{\boldsymbol{\beta}})$, in (1.12) and (1.13), we performed a simulation study. Based on $\boldsymbol{\beta}_t$ we simulated $1,000$ datasets with choices y_{jsn} from $N = 50$ respondents with respect to each Bayesian optimal design by drawing for each choice set s of the design and for each respondent n a random number v_{sn} from the uniform distribution $\mathcal{U}[0,1]$. These random numbers represent cumulative probabilities which we compared with the true logit probabilities $\mathbf{p}_s(\boldsymbol{\beta}_t)$ of the design to assign values to y_{jsn} in the following way:

$$y_{jsn} = \begin{cases} 1 & \text{if } v_{sn} \in \left]\sum_{t=1}^{j-1} p_{ts}(\boldsymbol{\beta}_t), \sum_{t=1}^{j} p_{ts}(\boldsymbol{\beta}_t)\right], \\ 0 & \text{otherwise.} \end{cases} \tag{1.14}$$

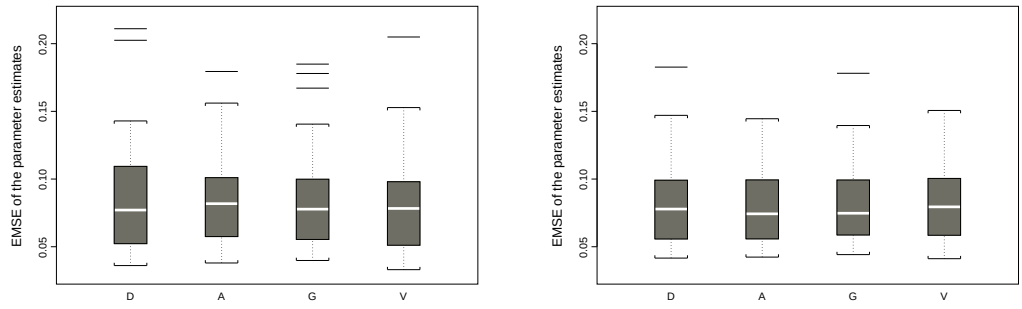
We obtained $1,000$ estimates $\hat{\boldsymbol{\beta}}$ by substituting the matrix $\mathbf{Y}$ for each dataset in the log-likelihood function (1.3) that we maximized with respect to $\boldsymbol{\beta}$. Because the $EMSE$ measures depend on a true parameter $\boldsymbol{\beta}_t$, we repeated their computation 50 times, each time for a different true parameter. Each computation for another $\boldsymbol{\beta}_t$ is called a replication.

We summarize the results of the 50 replications of the $EMSE_{\hat{\boldsymbol{\beta}}}$ in Panel a of Table 1.7 and in Figure 1.1. Using percentage values, Table 1.7, Panel a, depicts the number of replications for which the two- and three-alternative $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs have the lowest $EMSE_{\hat{\boldsymbol{\beta}}}$ value. The values themselves are collected in box plots in Figure 1.1. The white line in each of the boxes is the median. Overall, for the two design classes, it appears that there is no salient optimality criterion that leads to the most accurate estimates. From the box plots, we observe that the estimation performances of the design criteria are comparable. The median $EMSE_{\hat{\boldsymbol{\beta}}}$ values and the average $EMSE_{\hat{\boldsymbol{\beta}}}$ values, which are practically identical to the medians but not shown in the box plots, are equal across the different optimality criteria. Furthermore, Table 1.7, Panel a, indicates that the $\mathcal{G}_B$-optimality criterion has the smallest number of replications with the lowest $EMSE_{\hat{\boldsymbol{\beta}}}$ value. With regard to estimation differences between the two- and three-alternative designs, the box plots reveal that occasionally, larger $EMSE_{\hat{\boldsymbol{\beta}}}$ values are obtained for the two-alternative designs than for the three-alternative designs. Therefore, the parameter estimates produced by the three-alternative designs are somewhat more accurate.

Table 1.7: Percentage of replications with the lowest values for the $EMSE_{\hat{\boldsymbol{\beta}}}$ and $EMSE_{\hat{\mathbf{p}}_c}$ among the two- and three-alternative Bayesian optimal designs.

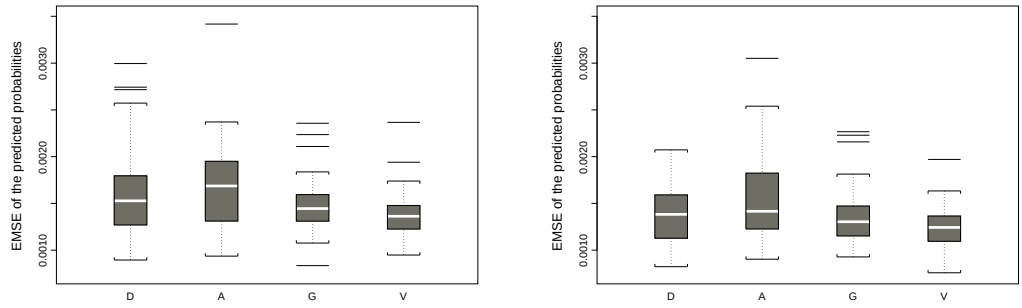
Optimal design	a) Rep. with lowest $EMSE_{\hat{\boldsymbol{\beta}}}$		b) Rep. with lowest $EMSE_{\hat{\mathbf{p}}_c}$	
	# Alternatives		# Alternatives	
	2	3	2	3
$\mathcal{D}_B$	26%	32%	18%	14%
$\mathcal{A}_B$	16%	28%	6%	10%
$\mathcal{G}_B$	14%	18%	16%	22%
$\mathcal{V}_B$	44%	22%	60%	54%

We carried out an analogous study for the 50 replications of the $EMSE_{\hat{\mathbf{p}}_c}$. The number of replications with the lowest $EMSE_{\hat{\mathbf{p}}_c}$ value for each of the optimal designs appears in Panel b of Table 1.7. The box plots with the $EMSE_{\hat{\mathbf{p}}_c}$ values for the two- and three-alternative optimal designs appear in Figure 1.2. As is illustrated by the medians in the plots, the occurrence of larger $EMSE_{\hat{\mathbf{p}}_c}$ values for the two-alternative than for the three-alternative designs is more pronounced. As a result, predictions based on the three-alternative designs tend to be more precise. With respect to the prediction performances of the optimality criteria, there are no real surprises. Table 1.7, Panel b, and the box plots palpably point toward the $\mathcal{V}_B$-optimality criterion as the criterion that provides the most precise predictions. The $\mathcal{G}_B$-optimality criterion is second best, followed by the $\mathcal{D}_B$-optimality criterion and the $\mathcal{A}_B$-optimality criterion. From the box plots, it is also apparent that the predictive capabilities of the customarily used $\mathcal{D}_B$-optimal designs do not differ that much from those of the $\mathcal{V}_B$- and $\mathcal{G}_B$-optimal designs, which are developed particularly for predictive purposes. Therefore, the $\mathcal{D}_B$-optimal designs seem to perform reasonably well in this respect.



(a) Two-alternative designs (b) Three-alternative designs

Figure 1.1: Distributions of the $EMSE_{\hat{\boldsymbol{\beta}}}$ obtained from 50 replications and computed for the two- and three-alternative $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs.



(a) Two-alternative designs (b) Three-alternative designs

Figure 1.2: Distributions of the $EMSE_{\hat{\mathbf{p}}_c}$ obtained from 50 replications and computed for the two- and three-alternative $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs.

1.4.2 Some computational aspects

We now embark on the account of some computational aspects of the two- and three-alternative $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs. We consecutively discuss the computing times to generate the designs and the computational effectiveness of the design criteria.

Computing time

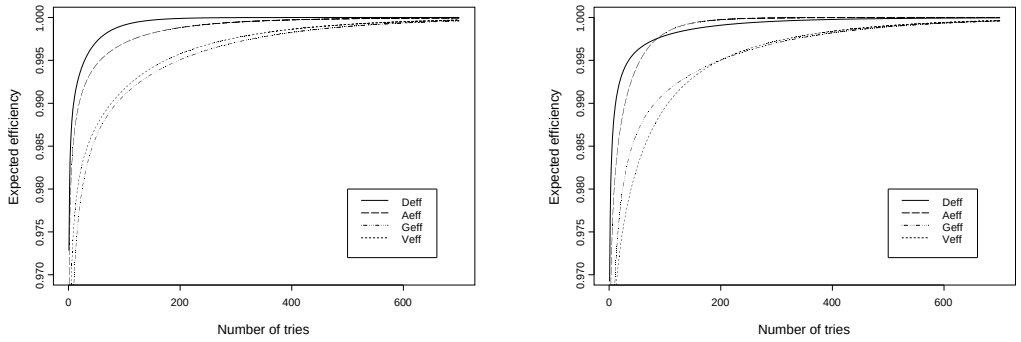
Table 1.8 reports computing times for one try of the modified Fedorov algorithm to produce the two- and three-alternative $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs. We generated the designs with the SAS 8.02 procedure IML. We obtained the times using a Dell personal computer with a 1.80 GHz Intel Processor and 256 MB RAM. Overtly, the computing times for the $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs are much longer than those for the $\mathcal{D}_B$- and $\mathcal{A}_B$-optimal designs. This is because of the numerous prediction variances that need to be computed when evaluating a design by means of the $\mathcal{G}_B$- or $\mathcal{V}_B$-optimality criterion. Furthermore, the number of prediction variances to derive is proportional to the design region χ that is eight times larger for the three-alternative design class than for the two-alternative class. This explains why it takes much more time to construct $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs with three alternatives than with two alternatives.

Table 1.8: Computing times for one try of the modified Fedorov algorithm to generate the two- and three-alternative Bayesian optimal designs. The times are expressed in hours:minutes.

Design criterion	# Alternatives	
	2	3
$\mathcal{D}_B$	00:05	00:05
$\mathcal{A}_B$	00:05	00:05
$\mathcal{G}_B$	02:30	11:00
$\mathcal{V}_B$	02:30	11:00

Computational effectiveness of the design criteria

The computational effectiveness of a Bayesian design criterion refers to the quality and the speed of the modified Fedorov algorithm in which this criterion is embedded. We compared the computational effectiveness of the Bayesian design criteria by means of the estimated expected efficiencies from several numbers of tries. The estimated expected efficiency of an optimal design produced by a number of tries, T, is defined as the efficiency to expect when T tries have been performed. We explain the calculation of the expected efficiency from T tries in Appendix D. For each of the two- and three-alternative $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs, we plotted the expected efficiencies against various numbers of tries. The plots appear in Figure 1.3. We obtained the highest expected efficiencies when we used the $\mathcal{D}_B$- and $\mathcal{A}_B$-optimality criteria. Applying the $\mathcal{G}_B$- and $\mathcal{V}_B$-optimality criteria requires more tries to reach a given efficiency. Consequently, the smallest number of tries is needed for calculating the $\mathcal{D}_B$- and $\mathcal{A}_B$-optimal designs. In addition, if the algorithm fails to find the $\mathcal{D}_B$- and $\mathcal{A}_B$-optimal designs, it still produces highly efficient designs.



(a) Two-alternative designs (b) Three-alternative designs

Figure 1.3: Estimated expected efficiencies of the two- and three-alternative $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs.

1.5 Conclusion

In this chapter, we incorporated the $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimality criteria in the modified Fedorov algorithm to generate two- and three-alternative Bayesian optimal choice designs containing the same number of profiles. We devoted special attention to the $\mathcal{G}_B$- and $\mathcal{V}_B$-optimality criteria which look for designs that produce precise predictions. After all, choice experiments are carried out to predict the future market share of related products or services as precisely as possible. We observed that the $\mathcal{V}_B$-optimal designs and, to a lesser extent, the $\mathcal{G}_B$-optimal designs are best suited for predictive purposes. The $\mathcal{D}_B$-optimal designs rank third in this aspect, but the differences in predictive ability compared with the $\mathcal{V}_B$- and $\mathcal{G}_B$-optimal designs are rather small. Furthermore, the three-alternative optimal designs lead to better predictions than the two-alternative designs. The three-alternative optimal designs also yield the most accurate parameter estimates, but there is no real difference in estimation performance between the distinct optimality criteria.

However, the computation of the $\mathcal{V}_B$- and $\mathcal{G}_B$-optimal designs takes a long time, particularly those with three alternatives, and many tries are needed. The $\mathcal{D}_B$- and $\mathcal{A}_B$-optimal designs are much faster to compute. To speed up the computations, one can slightly reduce the number of prior parameters drawn from the prior distribution when evaluating a design. Nevertheless, in weighing the large computational efforts against the small improvements in predictive ability of the $\mathcal{V}_B$- and $\mathcal{G}_B$-optimal designs, it seems preferable to retain the use of the $\mathcal{D}_B$-optimality criterion to build optimal choice designs. Moreover, as a rule of thumb, we cogently argue that if more than three attributes with more than two levels each are involved in the design optimization, the use of the $\mathcal{V}_B$- and $\mathcal{G}_B$-optimality criteria in combination with the modified Fedorov algorithm is no longer practically feasible. Drawing on the $\mathcal{V}_B$- and $\mathcal{G}_B$-optimality criteria to deal with large problem situations awaits the exploration of computationally more efficient algorithms. Finally, the Bayesian optimal designs are characterized by some level overlap and are not maximum utility balanced.

Appendix A. Derivation of the information matrix

The information matrix $\mathbf{M}(\mathbf{X}, \boldsymbol{\beta})$ in (1.4) can be estimated by $-\mathbf{H}(\mathbf{X}, \boldsymbol{\beta})$, where $\mathbf{H}(\mathbf{X}, \boldsymbol{\beta})$ is the Hessian matrix or the matrix of second-order derivatives of the log-likelihood function (1.3) with respect to $\boldsymbol{\beta}$. We begin by calculating the first-order derivative of the log-likelihood function given $N = 1$ respondent:

$$\begin{aligned}
\frac{\partial LL(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} =& \frac{\partial}{\partial \boldsymbol{\beta}} \sum_{s=1}^{S} \sum_{j=1}^{J} y_{js} \ln\left(p_{js}(\mathbf{X}_s, \boldsymbol{\beta})\right), \\
=& \sum_{s=1}^{S} \sum_{j=1}^{J} y_{js} \frac{\sum_{t=1}^{J} e^{\mathbf{x}'_{ts}\boldsymbol{\beta}}}{e^{\mathbf{x}'_{js}\boldsymbol{\beta}}} \left(\frac{e^{\mathbf{x}'_{js}\boldsymbol{\beta}} \mathbf{x}_{js} \sum_{t=1}^{J} e^{\mathbf{x}'_{ts}\boldsymbol{\beta}}}{\left(\sum_{t=1}^{J} e^{\mathbf{x}'_{ts}\boldsymbol{\beta}}\right)^2} \right) \\
& - \sum_{s=1}^{S} \sum_{j=1}^{J} y_{js} \frac{\sum_{t=1}^{J} e^{\mathbf{x}'_{ts}\boldsymbol{\beta}}}{e^{\mathbf{x}'_{js}\boldsymbol{\beta}}} \left(\frac{e^{\mathbf{x}'_{js}\boldsymbol{\beta}} \left(\sum_{t=1}^{J} e^{\mathbf{x}'_{ts}\boldsymbol{\beta}} \mathbf{x}_{ts}\right)}{\left(\sum_{t=1}^{J} e^{\mathbf{x}'_{ts}\boldsymbol{\beta}}\right)^2} \right), \\
=& \sum_{s=1}^{S} \sum_{j=1}^{J} y_{js} \mathbf{x}_{js} - \sum_{s=1}^{S} \sum_{j=1}^{J} y_{js} \frac{\sum_{t=1}^{J} e^{\mathbf{x}'_{ts}\boldsymbol{\beta}} \mathbf{x}_{ts}}{\sum_{t=1}^{J} e^{\mathbf{x}'_{ts}\boldsymbol{\beta}}}.
\end{aligned} \tag{A1}$$

The Hessian matrix $\mathbf{H}(\mathbf{X}, \boldsymbol{\beta})$ given $N = 1$ respondent then equals

$$\begin{aligned}
\frac{\partial^2 LL(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} =& \frac{\partial}{\partial \boldsymbol{\beta}'} \left(-\sum_{s=1}^{S} \sum_{j=1}^{J} y_{js} \frac{\sum_{t=1}^{J} e^{\mathbf{x}'_{ts}\boldsymbol{\beta}} \mathbf{x}_{ts}}{\sum_{t=1}^{J} e^{\mathbf{x}'_{ts}\boldsymbol{\beta}}} \right), \\
=& -\sum_{s=1}^{S} \sum_{j=1}^{J} y_{js} \left(\frac{\sum_{t=1}^{J} e^{\mathbf{x}'_{ts}\boldsymbol{\beta}} \mathbf{x}_{ts} \mathbf{x}'_{ts} \sum_{t=1}^{J} e^{\mathbf{x}'_{ts}\boldsymbol{\beta}}}{\left(\sum_{t=1}^{J} e^{\mathbf{x}'_{ts}\boldsymbol{\beta}} \right)^2} \right) \\
&+ \sum_{s=1}^{S} \sum_{j=1}^{J} y_{js} \left(\frac{\sum_{t=1}^{J} e^{\mathbf{x}'_{ts}\boldsymbol{\beta}} \mathbf{x}_{ts} \sum_{t=1}^{J} e^{\mathbf{x}'_{ts}\boldsymbol{\beta}} \mathbf{x}'_{ts}}{\left(\sum_{t=1}^{J} e^{\mathbf{x}'_{ts}\boldsymbol{\beta}} \right)^2} \right), \\
=& -\sum_{s=1}^{S} \sum_{j=1}^{J} y_{js} \left(\sum_{t=1}^{J} \frac{e^{\mathbf{x}'_{ts}\boldsymbol{\beta}}}{\sum_{v=1}^{J} e^{\mathbf{x}'_{vs}\boldsymbol{\beta}}} \mathbf{x}_{ts} \mathbf{x}'_{ts} \right) \\
&+ \sum_{s=1}^{S} \sum_{j=1}^{J} y_{js} \left(\sum_{t=1}^{J} \frac{e^{\mathbf{x}'_{ts}\boldsymbol{\beta}}}{\sum_{v=1}^{J} e^{\mathbf{x}'_{vs}\boldsymbol{\beta}}} \mathbf{x}_{ts} \right) \left(\sum_{t=1}^{J} \frac{e^{\mathbf{x}'_{ts}\boldsymbol{\beta}}}{\sum_{v=1}^{J} e^{\mathbf{x}'_{vs}\boldsymbol{\beta}}} \mathbf{x}'_{ts} \right), \\
=& -\sum_{s=1}^{S} \sum_{j=1}^{J} y_{js} \left(\sum_{t=1}^{J} p_{ts} \mathbf{x}_{ts} \mathbf{x}'_{ts} - \sum_{t=1}^{J} p_{ts} \mathbf{x}_{ts} \sum_{t=1}^{J} p_{ts} \mathbf{x}'_{ts} \right), \\
=& -\sum_{s=1}^{S} \left(\sum_{t=1}^{J} p_{ts} \mathbf{x}_{ts} \mathbf{x}'_{ts} - \sum_{t=1}^{J} p_{ts} \mathbf{x}_{ts} \sum_{t=1}^{J} p_{ts} \mathbf{x}'_{ts} \right), \\
=& -\sum_{s=1}^{S} \left(\mathbf{X}'_s \mathbf{P}_s \mathbf{X}_s - \mathbf{X}'_s \mathbf{p}_s \mathbf{p}'_s \mathbf{X}_s \right), \\
=& -\sum_{s=1}^{S} \mathbf{X}'_s \left(\mathbf{P}_s - \mathbf{p}_s \mathbf{p}'_s \right) \mathbf{X}_s,
\end{aligned} \tag{A2}$$

where $\mathbf{X}_s = [\mathbf{x}_{1s}, ..., \mathbf{x}_{Js}]'$, $\mathbf{p}_s = [p_{1s}, ..., p_{Js}]'$ and $\mathbf{P}_s = \text{diag}[p_{1s}, ..., p_{Js}]$. As such, the information matrix given N respondents is obtained as

$$\begin{aligned}
\mathbf{M}(\mathbf{X}, \boldsymbol{\beta}) =& N \left(-\frac{\partial^2 LL(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'} \right), \\
=& N \sum_{s=1}^{S} \mathbf{X}'_s \left(\mathbf{P}_s - \mathbf{p}_s \mathbf{p}'_s \right) \mathbf{X}_s.
\end{aligned} \tag{A3}$$

Appendix B. Numerical example

We compute the $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-criterion values of a small design consisting of three choice sets with two alternatives each. The alternatives include two attributes: Attribute 1 has three levels and Attribute 2 has two levels. The design matrix can be composed either by assigning numerical values (i.e., 1, 2, 3, and so forth) to the attribute levels or by employing effects-type coding. However, because of the categorical nature of the explanatory variables, it is more common to work with the design matrix from effects-type coding. With effects-type coding, the three levels of Attribute 1 are coded as [1 0], [0 1] and [−1 − 1], and the two levels of Attribute 2 are coded as −1 and 1. The design matrix containing numerical values, $\mathbf{X}^0$, and its companion in effects-type coding, $\mathbf{X}$, which we use in the computations, appear as follows:

$$\mathbf{X}^0 = \left[\begin{array}{cc} 1 & 2 \\ 2 & 1 \\ \hline 2 & 2 \\ 3 & 1 \\ \hline 3 & 2 \\ 1 & 1 \end{array}\right], \quad \mathbf{X} = \left[\begin{array}{ccc} 1 & 0 & 1 \\ 0 & 1 & -1 \\ \hline 0 & 1 & 1 \\ -1 & -1 & -1 \\ \hline -1 & -1 & 1 \\ 1 & 0 & -1 \end{array}\right]. \tag{B1}$$

The three choice sets are separated by horizontal lines. Each row specifies an alternative and the column dimension of the design matrix $\mathbf{X}$ corresponds to the number of parameters k. Here, $k = 3$. We compute the Bayesian criterion values as we do in (1.5), (1.6), (1.7) and (1.10). For the sake of illustration, we use only three prior parameters $\boldsymbol{\beta}^r = [\beta_{11}^r, \beta_{12}^r, \beta_2^r]'$, $r = 1, 2, 3$, randomly drawn from $\pi(\boldsymbol{\beta}) = \mathcal{N}(\boldsymbol{\beta}|\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$, where $\boldsymbol{\beta}_0 = [-1, 0, -1]'$ and $\boldsymbol{\Sigma}_0 = \mathbf{I}_3$. For each of these parameters, we compute the local $\mathcal{D}_P^r$-, $\mathcal{A}_P^r$-, $\mathcal{G}_P^r$- and $\mathcal{V}_P^r$-criterion values and subsequently average them to obtain the Bayesian values.

We compute the information matrix $\mathbf{M}(\mathbf{X}, \boldsymbol{\beta}^r)$ as we do in (1.4) by taking $N = 1$ so that $\mathbf{M}(\mathbf{X}, \boldsymbol{\beta}^r) = \sum_{s=1}^{3} \mathbf{X}_s'(\mathbf{P}_s - \mathbf{p}_s\mathbf{p}_s')\mathbf{X}_s$ with choice sets given by

$$\mathbf{X}_1 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}, \mathbf{X}_2 = \begin{bmatrix} 0 & 1 & 1 \\ -1 & -1 & -1 \end{bmatrix}, \mathbf{X}_3 = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix}, \tag{B2}$$

where $\mathbf{p}_s$ is the vector of probabilities in choice set s and $\mathbf{P}_s$ is the corresponding diagonal matrix. As a first draw, we have $\boldsymbol{\beta}^1 = [0.805, -0.080, -0.603]'$. Using the multinomial logit model (1.2), we obtain the following for choice set $s = 1$:

$$\mathbf{p}_1 = \begin{bmatrix} 0.420 \\ 0.580 \end{bmatrix}, \mathbf{P}_1 - \mathbf{p}_1\mathbf{p}_1' = 0.244 \times \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \text{ and} \tag{B3}$$

$$\mathbf{X}_1'(\mathbf{P}_1 - \mathbf{p}_1\mathbf{p}_1')\mathbf{X}_1 = \begin{bmatrix} 0.244 & -0.244 & 0.487 \\ -0.244 & 0.244 & -0.487 \\ 0.487 & -0.487 & 0.975 \end{bmatrix}. \tag{B4}$$

Repeating the computations for choice sets 2 and 3 and summing the three matrices yields the information matrix pertaining to $\boldsymbol{\beta}^1$:

$$\mathbf{M}(\mathbf{X}, \boldsymbol{\beta}^1) = \begin{bmatrix} 0.703 & 0.333 & 0.721 \\ 0.333 & 1.226 & 0.324 \\ 0.721 & 0.324 & 2.128 \end{bmatrix}. \tag{B5}$$

The local $\mathcal{D}_P^1$-criterion value then becomes

$$\mathcal{D}_P^1 = \left\{\det(\mathbf{M}^{-1}(\mathbf{X}, \boldsymbol{\beta}^1))\right\}^{1/3} = 0.986, \tag{B6}$$

and the local $\mathcal{A}_P^1$-criterion value becomes

$$\mathcal{A}_P^1 = \text{tr}(\mathbf{M}^{-1}(\mathbf{X}, \boldsymbol{\beta}^1)) = 4.057. \tag{B7}$$

To obtain the local $\mathcal{G}_P^1$- and $\mathcal{V}_P^1$-criterion values, we compute the prediction variances over the design region χ that consists of all possible choice sets of size two. For our small example, there are $3 \times 2 = 6$ candidate profiles so that χ comprises $Q = \binom{6}{2} = 15$ choice sets, or 30 profiles. For each of these profiles, we compute the $\mathbf{c}$ vector according to (1.9):

$$\chi = \begin{array}{|c|} \hline \mathbf{x}_{11} \\ \mathbf{x}_{21} \\ \hline \vdots \\ \hline \mathbf{x}_{1,15} \\ \mathbf{x}_{2,15} \\ \hline \end{array} = \begin{array}{|ccc|} \hline 1 & 0 & -1 \\ 1 & 0 & 1 \\ \hline \multicolumn{3}{|c|}{\vdots} \\ \hline -1 & -1 & -1 \\ -1 & -1 & 1 \\ \hline \end{array} \begin{array}{ccc} \rightarrow & p_{11} = 0.770 & \rightarrow \\ \rightarrow & p_{21} = 0.230 & \rightarrow \\ & & \\ \rightarrow & p_{1,15} = 0.770 & \rightarrow \\ \rightarrow & p_{2,15} = 0.230 & \rightarrow \end{array} \tag{B8}$$

$$\begin{aligned} \mathbf{c}(\mathbf{x}_{11}) &= [0, 0, -0.354]', \\ \mathbf{c}(\mathbf{x}_{21}) &= [0, 0, 0.354]', \\ & \\ \mathbf{c}(\mathbf{x}_{1,15}) &= [0, 0, -0.354]', \\ \mathbf{c}(\mathbf{x}_{2,15}) &= [0, 0, 0.354]'. \end{aligned}$$

Then, we derive the local $\mathcal{G}_P^1$- and $\mathcal{V}_P^1$-criterion values as

$$\mathcal{G}_P^1 = \max \begin{pmatrix} \mathbf{c}'(\mathbf{x}_{11})\mathbf{M}^{-1}(\mathbf{X}, \boldsymbol{\beta}^1)\mathbf{c}(\mathbf{x}_{11}) = 0.090 \\ \mathbf{c}'(\mathbf{x}_{21})\mathbf{M}^{-1}(\mathbf{X}, \boldsymbol{\beta}^1)\mathbf{c}(\mathbf{x}_{21}) = 0.090 \\ \vdots \\ \mathbf{c}'(\mathbf{x}_{1,15})\mathbf{M}^{-1}(\mathbf{X}, \boldsymbol{\beta}^1)\mathbf{c}(\mathbf{x}_{1,15}) = 0.090 \\ \mathbf{c}'(\mathbf{x}_{2,15})\mathbf{M}^{-1}(\mathbf{X}, \boldsymbol{\beta}^1)\mathbf{c}(\mathbf{x}_{2,15}) = 0.090 \end{pmatrix} = 0.338 \text{ and} \tag{B9}$$

$$\mathcal{V}_P^1 = \text{avg} \begin{pmatrix} \mathbf{c}'(\mathbf{x}_{11})\mathbf{M}^{-1}(\mathbf{X}, \boldsymbol{\beta}^1)\mathbf{c}(\mathbf{x}_{11}) = 0.090 \\ \mathbf{c}'(\mathbf{x}_{21})\mathbf{M}^{-1}(\mathbf{X}, \boldsymbol{\beta}^1)\mathbf{c}(\mathbf{x}_{21}) = 0.090 \\ \vdots \\ \mathbf{c}'(\mathbf{x}_{1,15})\mathbf{M}^{-1}(\mathbf{X}, \boldsymbol{\beta}^1)\mathbf{c}(\mathbf{x}_{1,15}) = 0.090 \\ \mathbf{c}'(\mathbf{x}_{2,15})\mathbf{M}^{-1}(\mathbf{X}, \boldsymbol{\beta}^1)\mathbf{c}(\mathbf{x}_{2,15}) = 0.090 \end{pmatrix} = 0.170. \tag{B10}$$

Similar computations for a second draw, $\boldsymbol{\beta}^2 = [-2.083, 2.238, -1.624]'$, yield

$$\begin{aligned} \mathcal{D}_P^2 &= 5.562, \\ \mathcal{A}_P^2 &= 486.821, \\ \mathcal{G}_P^2 &= 7.950, \\ \mathcal{V}_P^2 &= 1.163, \end{aligned} \tag{B11}$$

and for a third draw, $\boldsymbol{\beta}^3 = [-0.486, -0.087, -1.594]'$, we obtain

$$\begin{aligned} \mathcal{D}_P^3 &= 3.930, \\ \mathcal{A}_P^3 &= 17.804, \\ \mathcal{G}_P^3 &= 1.371, \\ \mathcal{V}_P^3 &= 0.462. \end{aligned} \tag{B12}$$

Finally, we average the local criterion values over the three draws to obtain the Bayesian values:

$$\begin{aligned} \mathcal{D}_B(\mathbf{X}) &= 3.493, \\ \mathcal{A}_B(\mathbf{X}) &= 169.561, \\ \mathcal{G}_B(\mathbf{X}) &= 3.219, \\ \mathcal{V}_B(\mathbf{X}) &= 0.598. \end{aligned} \tag{B13}$$

Appendix C. Minimum and maximum utility-balanced designs

Table C1: Bayesian minimum and maximum utility-balanced designs for the $3^2 \times 2/2/12$ example and their efficiencies with respect to the different optimality criteria.

		Min UB			Max UB		
Choice set	Alt	Attr 1	Attr 2	Attr 3	Attr 1	Attr 2	Attr 3
1	I	1	1	2	3	1	1
	II	3	3	1	3	2	1
2	I	1	1	2	3	1	2
	II	3	2	1	3	2	2
3	I	1	1	2	2	1	1
	II	2	3	1	2	2	1
4	I	1	2	2	2	1	2
	II	3	3	1	2	2	2
5	I	2	1	2	1	1	1
	II	3	3	1	1	2	1
6	I	1	1	2	1	1	2
	II	2	2	1	1	2	2
7	I	1	1	2	1	3	1
	II	3	1	1	2	3	1
8	I	1	2	2	1	3	2
	II	3	2	1	2	3	2
9	I	1	3	2	1	1	1
	II	3	3	1	2	1	1
10	I	1	2	2	2	3	1
	II	2	3	1	3	3	1
11	I	1	1	2	3	2	1
	II	3	3	2	3	3	1
12	I	1	1	1	1	3	1
	II	3	3	1	2	3	2

$\mathcal{D}_B$-efficiency	0.43%	37.92%
$\mathcal{A}_B$-efficiency	0.15%	18.49%
$\mathcal{G}_B$-efficiency	0.08%	8.50%
$\mathcal{V}_B$-efficiency	0.15%	15.25%

Table C2: Bayesian minimum and maximum utility-balanced designs for the $3^2 \times 2/3/8$ example and their efficiencies with respect to the different optimality criteria.

		Min UB			Max UB		
Choice set	Alt	Attr			Attr		
		1	2	3	1	2	3
1	I	1	1	2	1	2	1
	II	1	2	2	2	1	1
	III	3	3	1	2	2	1
2	I	1	1	2	1	2	2
	II	2	1	2	2	1	2
	III	3	3	1	2	2	2
3	I	1	1	2	1	1	1
	II	1	3	2	1	2	1
	III	3	3	1	2	1	1
4	I	1	1	2	1	1	2
	II	2	2	2	1	2	2
	III	3	3	1	2	1	2
5	I	1	2	2	2	2	1
	II	2	1	2	3	1	1
	III	3	3	1	3	2	1
6	I	1	1	2	2	2	2
	II	1	1	1	3	1	2
	III	3	3	1	3	2	2
7	I	1	1	2	1	3	2
	II	3	1	2	2	2	2
	III	3	3	1	2	3	2
8	I	1	2	2	3	1	1
	II	1	3	2	3	2	2
	III	3	3	1	3	2	1

$\mathcal{D}_B$-efficiency	0.24%	40.01%
$\mathcal{A}_B$-efficiency	0.07%	12.01%
$\mathcal{G}_B$-efficiency	0.04%	5.71%
$\mathcal{V}_B$-efficiency	0.08%	11.44%

Appendix D. Estimated expected efficiency

If T refers to the number of tries for the algorithm, the efficiency E_t of a design $\mathbf{X}_t$, $t = 1, ..., T$, generated by try t of the algorithm is given by

$$E_t = \frac{\mathcal{B}(\mathbf{X}^*)}{\mathcal{B}(\mathbf{X}_t)}, \tag{D1}$$

where $\mathcal{B}$ represents the $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- or $\mathcal{V}_B$-criterion value of a design and $\mathbf{X}^*$ is the optimal design according to that criterion.

Assume that for a large number of tries, T, we obtain G distinct designs $\mathbf{X}_1, ..., \mathbf{X}_G$, with efficiencies $E_1 > ... > E_G$ in terms of a particular optimality criterion. As such, $\mathbf{X}_1$ is the best design and an estimate of the probability of finding $\mathbf{X}_1$ in T tries, say π_1, is given by the number of times $\mathbf{X}_1$ is found divided by T. Correspondingly, if π_2, ..., π_G refer to the probabilities of finding $\mathbf{X}_2, ..., \mathbf{X}_G$ in T tries, the estimated expected efficiency from T tries is given by

$$\text{E(efficiency)} = \sum_{i=1}^{G-1} \left\{ \left(\sum_{j=i}^{G} \pi_j \right)^T - \left(\sum_{j=i+1}^{G} \pi_j \right)^T \right\} E_i + \pi_G^T E_G. \tag{D2}$$

The mathematical derivation underlying this expression can be retrieved in the work of Trinca and Gilmour (2000) who introduced the estimated expected efficiency in the context of block designs.

Chapter 2

An efficient algorithm for constructing choice designs

This chapter has been submitted as

☞ Kessels, R., Jones, B., Goos, P. and Vandebroek M. (2006). An efficient algorithm for constructing Bayesian optimal choice designs.

Abstract

Recently, Kessels, Goos and Vandebroek (2006) developed a way to produce Bayesian $\mathcal{G}$- and $\mathcal{V}$-optimal designs for the multinomial logit model. These designs allow for precise response predictions which is the goal of conjoint choice experiments. The authors showed that the $\mathcal{G}$- and $\mathcal{V}$-optimality criteria outperform the $\mathcal{D}$- and $\mathcal{A}$-optimality criteria for prediction. However, their $\mathcal{G}$- and $\mathcal{V}$-optimal design algorithm is computationally intensive, which is a barrier to its use in practice. In this chapter, we present an efficient algorithm for calculating Bayesian optimal designs by means of the different criteria. Particularly, the speed of computation for the $\mathcal{V}$-optimality criterion has improved dramatically. The new algorithm makes it possible to use Bayesian $\mathcal{D}$-, $\mathcal{A}$-, $\mathcal{G}$- and $\mathcal{V}$-optimal designs that are tailored to individual respondents in computerized conjoint choice studies.

2.1 Introduction

Conjoint choice experiments or more concisely, choice experiments, are widely used in marketing to measure how the attributes of a product or service jointly affect consumer preferences. In a choice experiment, a product or service is characterized by a combination of attribute levels called a profile or an alternative. Respondents then choose one from a group of profiles called a choice set. They repeat this task for several other choice sets presented to them. All submitted choice sets make up the experimental design. The aim of a choice experiment is to estimate the importance of each attribute and its levels based on the respondents' preferences. The estimates are then exploited to mimic real marketplace choices by making predictions about consumers' future purchasing behavior.

Designing an efficient choice experiment involves selecting those choice sets that result in an accurately estimated model providing precise predictions. Kessels, Goos and Vandebroek (2006) compared four different design criteria based on the multinomial logit model to reach this goal. They studied the predictive performance of the $\mathcal{D}$- and $\mathcal{A}$-optimality criteria versus the $\mathcal{G}$- and $\mathcal{V}$-optimality criteria. Special attention was paid to the $\mathcal{G}$- and $\mathcal{V}$-optimality criteria which aim at making precise predictions. The authors were the first to work out these criteria for the multinomial logit model. On the other hand, the $\mathcal{D}$- and $\mathcal{A}$-optimality criteria focus on accurate estimates. Until now, the $\mathcal{D}$-optimality criterion has been most often employed to construct efficient choice designs (see Huber and Zwerina 1996; Sándor and Wedel 2001).

Because the multinomial logit model is nonlinear in the parameters, the computation of the optimality criteria depends on the unknown parameter vector. To solve this problem, Kessels, Goos and Vandebroek (2006) adopted a Bayesian design procedure as proposed by Sándor and Wedel (2001). Following these authors, they approximated the design criteria using a Monte Carlo sample from a multivariate normal prior parameter distribution. Monte Carlo sampling involves taking a large number of random draws from a probability distribution as a surrogate for that distribution. Like Sándor and Wedel (2001), Kessels, Goos and Vandebroek (2006) used $1,000$ random draws. The four optimality criteria in the Bayesian context are labelled the $\mathcal{D}_B$-, $\mathcal{A}_B$-,

$\mathcal{G}_B$- and $\mathcal{V}_B$-optimality criteria. Kessels, Goos and Vandebroek (2006) implemented these criteria in a modified Fedorov algorithm (Cook and Nachtsheim 1980; Fedorov 1972) to construct $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs. We refer to their complete algorithm as the Monte Carlo modified Fedorov algorithm (MCMF).

Kessels, Goos and Vandebroek (2006) showed that the $\mathcal{G}_B$- and $\mathcal{V}_B$-optimality criteria outperform the $\mathcal{D}_B$- and $\mathcal{A}_B$-optimality criteria in terms of prediction accuracy. They warn, however, that the computation of $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs is substantially more demanding than the search for $\mathcal{D}_B$- and $\mathcal{A}_B$-optimal designs. The long computing times resulting from MCMF make the $\mathcal{G}_B$- and $\mathcal{V}_B$-optimality criteria impractical to use. Also, the computational burden implies that the application of the $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimality criteria to computerized conjoint choice studies is limited. Ideally, computerized conjoint studies use choice designs that are tailored to the individual respondents so that maximum information is obtained on the individuals' preferences and thus on the heterogeneity between subjects.

The goal of this chapter is to present a novel design construction algorithm that is much faster than MCMF employed by Kessels, Goos and Vandebroek (2006). The speed of the new algorithm allows the $\mathcal{G}_B$- and $\mathcal{V}_B$-optimality criteria to be used in practice and it also opens the perspective of applying individualized Bayesian optimal choice designs in web-based conjoint studies.

The new algorithm has four key features. First, it uses an update formula to economically calculate the change in any of the optimality criteria for two designs that differ only in one profile. In this way, the optimality criterion values do not need to be re-computed from scratch. Second, it involves a formula for the $\mathcal{V}_B$-optimality criterion so that its computation is even more efficient. Third, the algorithm is an adaptation of Meyer and Nachtsheim's (1995) coordinate-exchange algorithm which is much faster than the modified Fedorov algorithm. Lastly, it relies on a designed sample of only 20 prior parameters instead of the Monte Carlo sample of $1,000$ draws. However, the algorithm still checks the designs produced by each random start using the Monte Carlo sample. Because of this re-evaluation, the algorithm is called

the adaptive algorithm.

The outline of the remainder of the chapter is as follows. Section 2.2 reviews the $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimality criteria for the multinomial logit model. In Section 2.3, we present the adaptive algorithm as an alternative to MCMF for faster computation of the optimal designs for all four criteria. We use the design example from Kessels, Goos and Vandebroek (2006) for comparison purposes. Section 2.4 discusses the four key features of the adaptive algorithm and Section 2.5 considers a more challenging scenario made possible by the faster method. Section 2.6 summarizes the results and suggests some opportunities for further research.

2.2 Design criteria for the multinomial logit

To present our improved design construction approach, we start with an overview of the different design criteria for the multinomial logit model. The model draws on a choice design matrix $\mathbf{X} = [\mathbf{x}'_{js}]_{j=1,...,J;s=1,...,S}$, where $\mathbf{x}_{js}$ is a $k \times 1$ vector of the attribute levels of profile j in choice set s. A respondent's utility for that profile is modelled as $U_{js} = \mathbf{x}'_{js}\boldsymbol{\beta} + \varepsilon_{js}$, where $\boldsymbol{\beta}$ is a $k \times 1$ vector of parameters and ε_{js} is an i.i.d. extreme value error term. The multinomial logit probability a respondent chooses profile j in choice set s is $p_{js} = e^{\mathbf{x}'_{js}\boldsymbol{\beta}} / \sum_{t=1}^{J} e^{\mathbf{x}'_{ts}\boldsymbol{\beta}}$. The information matrix $\mathbf{M}$, which is the inverse of the variance-covariance matrix of the parameter estimators, is the sum of the information matrices of the S choice sets $\mathbf{M}_s$ as shown below:

$$\begin{aligned} \mathbf{M}(\mathbf{X}, \boldsymbol{\beta}) &= N \sum_{s=1}^{S} \mathbf{M}_s(\mathbf{X}_s, \boldsymbol{\beta}) \\ &= N \sum_{s=1}^{S} \mathbf{X}'_s(\mathbf{P}_s - \mathbf{p}_s\mathbf{p}'_s)\mathbf{X}_s, \end{aligned} \tag{2.1}$$

where $\mathbf{X}_s = [\mathbf{x}_{1s}, ..., \mathbf{x}_{Js}]'$, $\mathbf{p}_s = [p_{1s}, ..., p_{Js}]'$, $\mathbf{P}_s = \text{diag}[p_{1s}, ..., p_{Js}]$ and N is the number of respondents. Kessels, Goos and Vandebroek (2006) implemented different design criteria or functions of the information matrix (2.1) for constructing optimal choice designs. This task is complicated by the fact that the information on the parameters depends on the unknown

values of those parameters through the probabilities. Therefore, the authors adopted a Bayesian design strategy that integrates the design criteria over a prior parameter distribution $\pi(\boldsymbol{\beta})$. The multivariate normal distribution $N(\boldsymbol{\beta}|\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$ was chosen for this purpose.

The design criteria employed are the $\mathcal{D}$-, $\mathcal{A}$-, $\mathcal{G}$- and $\mathcal{V}$-optimality criteria. The $\mathcal{D}$- and $\mathcal{A}$-optimality criteria both are concerned with a precise estimation of the parameters $\boldsymbol{\beta}$ in the multinomial logit model. The $\mathcal{D}$-optimality criterion aims at designs that minimize the determinant of the variance-covariance matrix of the parameter estimators, while the $\mathcal{A}$-optimality criterion aims at designs that minimize the trace of the variance-covariance matrix. The Bayesian $\mathcal{D}$-optimality criterion is

$$\mathcal{D}_B = \int_{\mathcal{R}^k} \left\{\det(\mathbf{M}^{-1}(\mathbf{X}, \boldsymbol{\beta}))\right\}^{1/k} \pi(\boldsymbol{\beta}) \mathrm{d}\boldsymbol{\beta}, \tag{2.2}$$

with the $\mathcal{D}_B$-optimal design minimizing (2.2). The $\mathcal{A}_B$-optimal design minimizes

$$\mathcal{A}_B = \int_{\mathcal{R}^k} \mathrm{tr}(\mathbf{M}^{-1}(\mathbf{X}, \boldsymbol{\beta}))\pi(\boldsymbol{\beta}) \mathrm{d}\boldsymbol{\beta}. \tag{2.3}$$

The $\mathcal{G}$- and $\mathcal{V}$-optimality criteria were developed to make precise response predictions. These criteria are important in this context since predicting consumer responses is the goal of choice experiments. The $\mathcal{G}$- and $\mathcal{V}$-optimality criteria for the multinomial logit model were first elaborated by Kessels, Goos and Vandebroek (2006). They are defined with respect to a design region χ consisting of all Q possible choice sets of size J that can be composed from the candidate profiles: $\chi = \{\{\mathbf{x}_{1q}, ..., \mathbf{x}_{Jq}\}|\ q = 1, ..., Q\}$. A $\mathcal{G}$-optimal design minimizes the maximum prediction variance over the design region χ, while a $\mathcal{V}$-optimality design minimizes the average prediction variance over this region. Formally, the $\mathcal{G}_B$-optimality criterion is

$$\begin{aligned} \mathcal{G}_B &= \int_{\mathcal{R}^k} \max_{\mathbf{x}_{jq} \in \chi} \mathrm{var}\{\hat{p}_{jq}(\mathbf{x}_{jq}, \boldsymbol{\beta})\}\pi(\boldsymbol{\beta}) \mathrm{d}\boldsymbol{\beta} \\ &= \int_{\mathcal{R}^k} \max_{\mathbf{x}_{jq} \in \chi} \mathbf{c}'(\mathbf{x}_{jq}) \mathbf{M}^{-1}(\mathbf{X}, \boldsymbol{\beta}) \mathbf{c}(\mathbf{x}_{jq}) \pi(\boldsymbol{\beta}) \mathrm{d}\boldsymbol{\beta}, \end{aligned} \tag{2.4}$$

where $\hat{p}_{jq}(\mathbf{x}_{jq}, \boldsymbol{\beta})$ denotes the predicted choice probability for $\mathbf{x}_{jq}$ and

$$\mathbf{c}(\mathbf{x}_{jq}) = \frac{\partial p_{jq}(\mathbf{x}_{jq}, \boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = p_{jq} \left(\mathbf{x}_{jq} - \sum_{t=1}^{J} p_{tq} \mathbf{x}_{tq} \right), \tag{2.5}$$

the partial derivative of the multinomial logit probability with respect to $\boldsymbol{\beta}$. The $\mathcal{V}_B$-optimality criterion is

$$\mathcal{V}_B = \int_{\mathcal{R}^k} \int_{\chi} \mathbf{c}'(\mathbf{x}_{jq}) \mathbf{M}^{-1}(\mathbf{X}, \boldsymbol{\beta}) \mathbf{c}(\mathbf{x}_{jq}) \mathrm{d}\mathbf{x}_{jq} \pi(\boldsymbol{\beta}) \mathrm{d}\boldsymbol{\beta} \tag{2.6}$$

with $\mathbf{c}(\mathbf{x}_{jq})$ given by (2.5).

2.3 The adaptive algorithm versus MCMF

We propose the adaptive algorithm for generating $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs instead of the Monte Carlo modified Fedorov algorithm (MCMF) employed by Kessels, Goos and Vandebroek (2006) (see Section 2.1). The adaptive algorithm is much faster than MCMF so that for a given computing time the resulting designs outperform the designs produced by MCMF.

We illustrate the better results from the adaptive algorithm versus MCMF using the design example of Kessels, Goos and Vandebroek (2006). These authors constructed $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs of two classes: $3^2 \times 2/2/12$ and $3^2 \times 2/3/8$. The design profiles in the two classes have a similar attribute structure with two attributes at three levels and one attribute at two levels. Hence, the sets of candidate profiles of the classes comprise the same $3^2 \times 2 = 18$ profiles. The designs of the first class consist of 12 choice sets of size two, while the designs of the second class consist of 8 choice sets of size three. So, the designs of both classes contain 24 profiles. Since we exploit this design example of 24 profiles to compare the adaptive algorithm with MCMF, we refer to it as the comparison example and label it $3^2 \times 2/24$.

Using effects-type coding (see Kessels, Goos and Vandebroek 2006), the number of elements, k, in the parameter vector is five. As prior parameter distribution, Kessels, Goos and Vandebroek (2006) proposed the multivariate normal distribution $\pi(\boldsymbol{\beta}) = \mathcal{N}(\boldsymbol{\beta}|\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$, with $\boldsymbol{\beta}_0 = [-1, 0, -1, 0, -1]'$ and $\boldsymbol{\Sigma}_0 = \mathbf{I}_5$. They approximated this distribution by drawing a Monte Carlo sample of $R = 1,000$ prior parameter values $\boldsymbol{\beta}^r$, $r = 1, ..., R$, from it. The

Bayesian optimal designs were then obtained from 200 tries or random starts of the modified Fedorov algorithm. This algorithm iteratively improves the starting design by exchanging its profiles with profiles from the candidate set. To compute the $\mathcal{G}_B$- and $\mathcal{V}_B$-optimality criteria for the two-alternative designs, the design region χ consists of $Q = \binom{18}{2} = 153$ choice sets, or 306 profiles, whereas for the three-alternative designs, it includes $Q = \binom{18}{3} = 816$ choice sets, or $2,448$ profiles.

Based on the same normal prior distribution we employed the adaptive algorithm to reproduce the $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs for the comparison example. Besides the two- and three-alternative designs, we also generated the four-alternative designs containing six choice sets. The design region χ in this case is quite extensive involving $Q = \binom{18}{4} = 3,060$ choice sets, or $12,240$ profiles. The optimal designs from the adaptive algorithm appear in Tables A1, A2 and A3 of the Appendix. In Table 2.1, we compared their criterion values with the criterion values from MCMF that we copied from the work of Kessels, Goos and Vandebroek (2006). As can be seen, the two-alternative $\mathcal{D}_B$-optimal designs from both algorithms are equivalent. However, in all the other cases with two and three alternatives, the designs generated with the adaptive algorithm outperform the designs generated with MCMF.

Table 2.1: $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-criterion values of the $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs for the comparison example $3^2 \times 2/24$ computed using the adaptive algorithm and the Monte Carlo modified Fedorov algorithm.

Optimal design	2 alternatives		3 alternatives		4 alternatives	
	Adaptive	MCMF	Adaptive	MCMF	Adaptive	MCMF
$\mathcal{D}_B$	0.73024	0.73024	0.75362	0.76617	0.86782	—
$\mathcal{A}_B$	6.55212	6.60563	5.97903	6.02261	6.57135	—
$\mathcal{G}_B$	0.49887	0.51997	0.51051	0.51843	0.60494	—
$\mathcal{V}_B$	0.07184	0.07219	0.06267	0.06285	0.05728	—

The best criterion values from the adaptive algorithm were the result of $1,000$ random starts rather than the 200 random starts utilized to obtain the best criterion values from MCMF. Because the adaptive algorithm is so much faster than MCMF, the extra random starts were still accomplished using far less computing time. The computing times for one try of the adaptive algorithm and MCMF appear in Tables 2.2a and 2.2b, respectively. We performed all computations in MATLAB 7 using a Dell personal computer with a 1.60 GHz Intel Processor and 2 GB RAM.

Tables 2.2a and 2.2b show the huge reductions in computing time using the adaptive algorithm. Particularly important are the reductions in computing time for the $\mathcal{G}_B$- and $\mathcal{V}_B$-optimality criteria. With the adaptive algorithm the construction of the $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs has become practically feasible. Even the four-alternative $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs were generated quickly, while their computation was not doable with MCMF. Notice also the faster running time for the $\mathcal{V}_B$-optimality criterion compared with the $\mathcal{G}_B$-optimality criterion. This is due to a computational short cut in the calculation of the $\mathcal{V}_B$-optimality criterion which we lay out in Section 2.4.2.

Table 2.2: Computing times for one try of the adaptive algorithm and the Monte Carlo modified Fedorov algorithm to generate the Bayesian optimal designs for the comparison example $3^2 \times 2/24$. The times are expressed in hours:minutes:seconds.

a) Adaptive algorithm

Design criterion	# Alternatives		
	2	3	4
$\mathcal{D}_B$	00:00:03	00:00:04	00:00:05
$\mathcal{A}_B$	00:00:03	00:00:04	00:00:05
$\mathcal{G}_B$	00:00:07	00:00:32	00:04:23
$\mathcal{V}_B$	00:00:03	00:00:05	00:00:08

b) Monte Carlo modified Fedorov

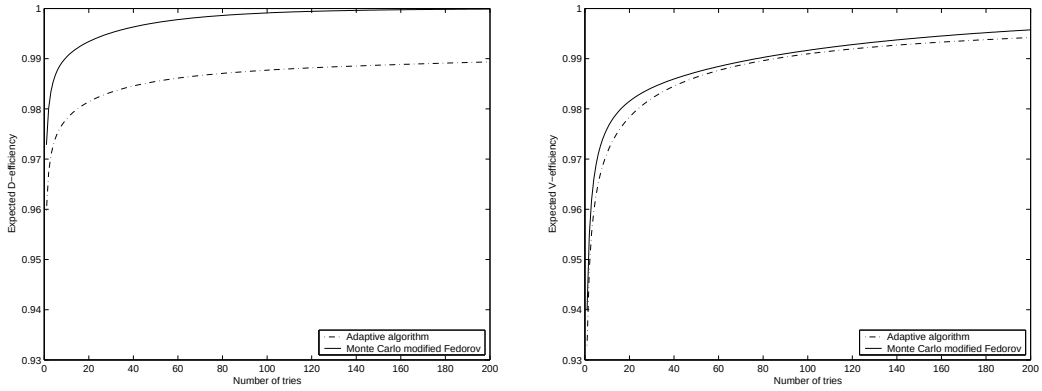
Design criterion	# Alternatives		
	2	3	4
$\mathcal{D}_B$	00:08:00	00:08:00	—
$\mathcal{A}_B$	00:08:00	00:08:00	—
$\mathcal{G}_B$	03:00:00	12:00:00	—
$\mathcal{V}_B$	03:00:00	12:00:00	—

Note that the adaptive algorithm is computationally less effective per number of tries than MCMF. This can be seen from the plots in Figure 2.1 in which we compare the estimated expected efficiencies against various numbers of tries of the adaptive algorithm and MCMF for computing the two-alternative

$\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs. These are the efficiencies to expect if a number of tries are performed with each of the algorithms. Details on the calculation of the expected efficiency from a number of tries can be found in the work of Kessels, Goos and Vandebroek (2006).

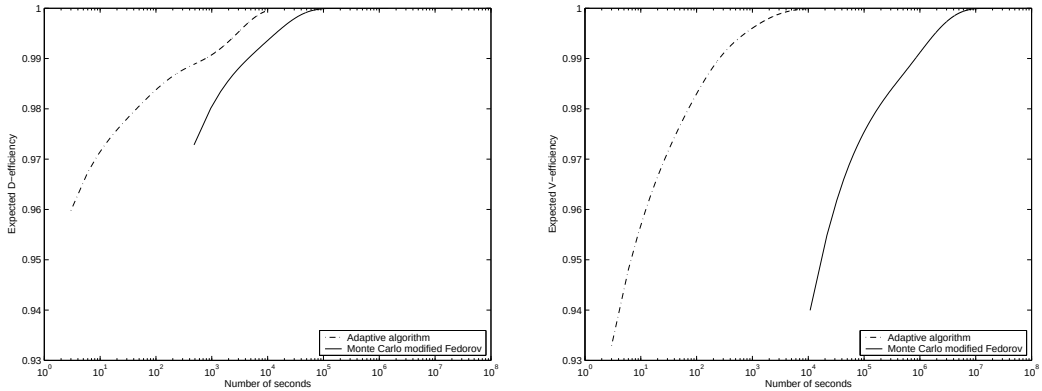
The plots for the two-alternative $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs in Figure 2.1 are also representative of the two-alternative $\mathcal{A}_B$- and $\mathcal{G}_B$-optimal designs, respectively. The plots for the three-alternative designs exhibit a similar pattern. From the plots, we observe that the differences in efficiency in favor of MCMF are smaller when a prediction-based design criterion is used instead of an estimation-based design criterion. This might be due to the fact that design optimization with the $\mathcal{G}_B$- and $\mathcal{V}_B$-optimality criteria is generally more difficult than with the $\mathcal{D}_B$- and $\mathcal{A}_B$-optimality criteria.

A more realistic comparison of the effectiveness of the adaptive algorithm versus MCMF appears in the plots of Figure 2.2. In these graphs, we plotted the estimated expected efficiencies of the two-alternative $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs against the number of seconds of computing time. We expressed the number of seconds on a log-scale. These plots provide compelling evidence of the practical value of the adaptive algorithm. The huge increase in speed created by the adaptive algorithm overtly leads to more efficient designs in a given amount of computing time. This is especially the case for the prediction-based design criteria as illustrated by the plot for the $\mathcal{V}_B$-efficiencies. Note however, that the bend in the plot for the $\mathcal{D}_B$-efficiencies reveals that the adaptive algorithm has a little difficulty making the final jump from 99% efficiency to 100% or global efficiency.



(a) Expected $\mathcal{D}_B$-efficiencies (b) Expected $\mathcal{V}_B$-efficiencies

Figure 2.1: Estimated expected efficiencies against various numbers of tries of the adaptive algorithm and the Monte Carlo modified Fedorov algorithm for computing the two-alternative $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs.



(a) Expected $\mathcal{D}_B$-efficiencies (b) Expected $\mathcal{V}_B$-efficiencies

Figure 2.2: Estimated expected efficiencies against various numbers of seconds of the adaptive algorithm and the Monte Carlo modified Fedorov algorithm for computing the two-alternative $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs.

2.4 Features of the adaptive algorithm

There are four features of the adaptive algorithm that result in increased speed compared with MCMF. They are:

1. updating the Cholesky decomposition of the information matrix,
2. an efficient computation of the $\mathcal{V}_B$-optimality criterion,
3. a coordinate-exchange algorithm,
4. a small designed sample of prior parameters.

The next sections discuss each of these in succession.

2.4.1 Updating the Cholesky decomposition of the information matrix

Updating the Cholesky decomposition of the information matrix is an economical way to compute the $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-criterion values of designs that differ only in one profile from another design. The Cholesky decomposition forms a symmetric positive definite matrix as an upper triangular matrix multiplied on the left by its transpose. The information matrix $\mathbf{M}$ is symmetric because the information matrices of the S choice sets $\mathbf{M}_s$ are symmetric. They are of the form $\mathbf{X}_s'\mathbf{C}_s\mathbf{X}_s$, where $\mathbf{C}_s = \mathbf{P}_s - \mathbf{p}_s\mathbf{p}_s'$ is symmetric. If $\mathbf{M}$ is positive definite, then its Cholesky decomposition is defined as

$$\mathbf{M} = \mathbf{L}'\mathbf{L}, \tag{2.7}$$

where $\mathbf{L}$ is an upper triangular matrix named the Cholesky factor.

In the adaptive algorithm, different designs are generated by changing only one attribute level of a single profile at a time (see Section 2.4.3). The starting design is denoted by $\mathbf{X}^s$. We compute the $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-criterion values of each of the designs as follows. For each prior parameter vector, we compute the information matrix $\mathbf{M}^s$ through (2.1) and derive its Cholesky factor $\mathbf{L}^s$. We then update the Cholesky factor after every profile change with low rank updates based on the work of Bennett (1965). Using the Cholesky factors the four criterion values for each design can be obtained

as shown below. In this way, we avoid re-computation of the information matrix through (2.1). For the comparison example $3^2 \times 2/24$, this procedure reduced the computing times by roughly a factor of three.

We now illustrate how the different design criteria rely on the Cholesky factor $\mathbf{L}$ of the information matrix $\mathbf{M}$. For any vector of coefficients, the $\mathcal{D}$-optimality criterion becomes

$$\begin{aligned}\mathcal{D} =&(\det(\mathbf{M}^{-1}))^{1/k} = 1/(\det(\mathbf{M}))^{1/k} = 1/(\det(\mathbf{L}')\det(\mathbf{L}))^{1/k} \\ =&1 \Big/ \left(\prod_{i=1}^{k} l_{ii}\right)^{2/k},\end{aligned} \tag{2.8}$$

where l_{ii} is the ith diagonal element of $\mathbf{L}$. Thus, to obtain the $\mathcal{D}_B$-criterion value of a design in which a profile has been changed, we do not need to re-compute the information matrix for every prior parameter vector. Only an update of the Cholesky factor is required.

To show the dependency of the $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimality criteria on the Cholesky factor $\mathbf{L}$, the Cholesky decomposition (2.7) has to be inverted. Denoting $\mathbf{L}^{-1}$ by $\mathbf{L}_{inv}$, the inverse is given by

$$\mathbf{M}^{-1} = (\mathbf{L}'\mathbf{L})^{-1} = \mathbf{L}_{inv}\mathbf{L}'_{inv}. \tag{2.9}$$

Because the Cholesky factor, $\mathbf{L}$, is triangular, inverting it is easier than inverting $\mathbf{M}$. Then, for any prior parameter vector, the $\mathcal{A}$-optimality criterion is

$$\mathcal{A} = \operatorname{tr}(\mathbf{M}^{-1}) = \operatorname{tr}(\mathbf{L}_{inv}\mathbf{L}'_{inv}) = \sum_{i=1}^{k}\sum_{j=i}^{k} m_{ij}^2, \tag{2.10}$$

where m_{ij} is the ijth element in $\mathbf{L}_{inv}$. So to obtain the $\mathcal{A}_B$-criterion value of a design in which a profile has been changed, we need to derive the new Cholesky factor for every prior parameter vector and take its inverse. This goes much faster than computing the new information matrix and inverting it.

In a similar manner, the $\mathcal{G}_B$- and $\mathcal{V}_B$-criterion values are obtained. The prediction variance of profile $\mathbf{x}_{jq} \in \chi$ is expressed as

$$\mathbf{c}'(\mathbf{x}_{jq})\mathbf{M}^{-1}\mathbf{c}(\mathbf{x}_{jq}) = \mathbf{c}'(\mathbf{x}_{jq})\mathbf{L}_{inv}\mathbf{L}'_{inv}\mathbf{c}(\mathbf{x}_{jq}). \tag{2.11}$$

Here, $\mathbf{c}(\mathbf{x}_{jq})$ does not depend on the design $\mathbf{X}$ and therefore only needs to be computed once for each prior parameter vector. The $\mathcal{G}_B$-criterion value is obtained by inserting (2.11) in (2.4). For the $\mathcal{V}_B$-optimality criterion, we performed some initial calculations that make its computation even more efficient. We describe these calculations in the next section.

2.4.2 Efficient computation of the $\mathcal{V}_B$-optimality criterion

In the adaptive algorithm, the $\mathcal{V}_B$-optimality criterion is implemented in an efficient way. For each prior vector of coefficients, it is possible to compute the average prediction variance without first computing the prediction variances for each profile $\mathbf{x}_{jq} \in \chi$ separately. A similar approach does not apply to the $\mathcal{G}_B$-optimality criterion since finding the worst prediction variance requires the computation of all variances.

To explain our method, we start from the prediction variance (2.11), but for the sake of clarity, we leave the implementation of the inverse of the Cholesky decomposition for the end. The prediction variance is naturally a scalar since $\mathbf{c}(\mathbf{x}_{jq})$ is a $k \times 1$ vector and $\mathbf{M}^{-1}$ a $k \times k$ matrix. The trace of a scalar is the scalar itself so that

$$\mathbf{c}'(\mathbf{x}_{jq})\mathbf{M}^{-1}\mathbf{c}(\mathbf{x}_{jq}) = \text{tr}(\mathbf{c}'(\mathbf{x}_{jq})\mathbf{M}^{-1}\mathbf{c}(\mathbf{x}_{jq})). \tag{2.12}$$

Now, $\text{tr}(\mathbf{ABC}) = \text{tr}(\mathbf{CAB})$ if $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ are matrices such that $\mathbf{ABC}$ is a square matrix and the matrix product $\mathbf{CAB}$ exists. This equality is known as the cyclic property of the trace. Since the prediction variance is a scalar and $\mathbf{c}(\mathbf{x}_{jq})\mathbf{c}'(\mathbf{x}_{jq})$ is a $k \times k$ matrix that conforms with $\mathbf{M}^{-1}$,

$$\text{tr}(\mathbf{c}'(\mathbf{x}_{jq})\mathbf{M}^{-1}\mathbf{c}(\mathbf{x}_{jq})) = \text{tr}(\mathbf{c}(\mathbf{x}_{jq})\mathbf{c}'(\mathbf{x}_{jq})\mathbf{M}^{-1}). \tag{2.13}$$

Let $\mathbf{W}_{jq} = \mathbf{c}(\mathbf{x}_{jq})\mathbf{c}'(\mathbf{x}_{jq})$. Because $\mathbf{c}(\mathbf{x}_{jq})$ does not depend on the design $\mathbf{X}$, $\mathbf{W}_{jq}$ is not a function of $\mathbf{X}$ either so that it only has to be computed once for each prior parameter vector. We now average the individual matrices $\mathbf{W}_{jq}$ over all profiles $\mathbf{x}_{jq} \in \chi$ and denote the subsequent matrix by $\mathbf{W}$:

$$\mathbf{W} = \frac{1}{JQ}\sum_{j=1}^{J}\sum_{q=1}^{Q}\mathbf{W}_{jq}. \tag{2.14}$$

The average prediction variance across all profiles $\mathbf{x}_{jq} \in \chi$ for a given prior parameter vector is then

$$\int_{\chi} \mathbf{c}'(\mathbf{x}_{jq})\mathbf{M}^{-1}\mathbf{c}(\mathbf{x}_{jq})\mathrm{d}\mathbf{x}_{jq} = \mathrm{tr}(\mathbf{W}\mathbf{M}^{-1}) \tag{2.15}$$

We refer to the work of Meyer and Nachtsheim (1995) for a similar expression of the $\mathcal{V}$-optimality criterion in the linear design setting. Finally, in terms of the inverse of the Cholesky decomposition of the information matrix (2.9), the average prediction variance is

$$\mathrm{tr}(\mathbf{W}\mathbf{M}^{-1}) = \mathrm{tr}(\mathbf{W}\mathbf{L}_{inv}\mathbf{L}'_{inv}). \tag{2.16}$$

So, to obtain the $\mathcal{V}_B$-optimality criterion, we have to compute $\mathbf{W}$ for each prior parameter vector only once. The set of $\mathbf{W}$ matrices can be re-used from one random start to the next.

2.4.3 Coordinate-exchange algorithm

The adaptive algorithm uses Meyer and Nachtsheim's (1995) coordinate-exchange algorithm to generate Bayesian optimal designs. As opposed to the modified Fedorov algorithm employed in Kessels, Goos and Vandebroek (2006), it allows the computation of choice designs with a large number of profiles, attributes and/or attribute levels in a reasonable amount of time. The coordinate-exchange algorithm can be seen as a greedy profile exchange algorithm. Whereas the modified Fedorov algorithm possibly changes every "coordinate" or attribute level of a profile, the coordinate-exchange algorithm only changes one coordinate. For each attribute level in the design, the coordinate-exchange algorithm tries all possible levels and chooses the level corresponding to the best value of the optimality criterion under consideration.

In contrast with the modified Fedorov algorithm, the coordinate-exchange algorithm is a candidate-set-free algorithm. That is, it does not require the specification of a set of candidate profiles. This aspect becomes more important when the candidate set is very large because of a large number of attributes and/or attribute levels. The coordinate-exchange algorithm is also

substantially faster than the modified Fedorov algorithm. It runs in polynomial time, while the modified Fedorov algorithm runs in exponential time. For the comparison example, this leads to roughly a factor of three speed increase of the coordinate-exchange algorithm over the modified Fedorov algorithm. For designs with more profiles, attributes and/or attribute levels, this increase in speed becomes more pronounced.

A small disadvantage of the coordinate-exchange algorithm compared with the modified Fedorov algorithm is that it generally takes more random starts to find a globally optimal design, especially when the $\mathcal{D}_B$- and $\mathcal{A}_B$-optimality criteria are utilized. The plots in Figure 2.1 with estimated expected $\mathcal{D}_B$- and $\mathcal{V}_B$-efficiencies for various numbers of tries illustrate this (see also Section 2.3). Nevertheless, if the global optimum is not reached, the coordinate-exchange algorithm still finds a very efficient design. Also, in terms of computing time, the coordinate-exchange algorithm may be more effective than the modified Fedorov algorithm. This is certainly the case for large, realistic design problems. Therefore, the lesser performance of the coordinate-exchange algorithm per number of tries can be disregarded.

The coordinate-exchange algorithm has also been applied by Kuhfeld and Tobias (2005) to generate $\mathcal{D}$-efficient factorial designs for large choice experiments based on a linear model. In their SAS `%MktEx` macro, the coordinate-exchange algorithm is incorporated together with the modified Fedorov algorithm and a large catalog of orthogonal arrays. If no orthogonal design meets the design problem and the modified Fedorov algorithm is impractical to use, then the coordinate-exchange algorithm is addressed. It may also be combined with simulated annealing.

2.4.4 Small designed sample of prior parameters

In this section, we present a new method to approximate the integral related to a multivariate normal prior $\pi(\boldsymbol{\beta}) = \mathcal{N}(\boldsymbol{\beta}|\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$ in the definitions of the Bayesian optimality criteria. The solution of the integral with respect to a multivariate normal prior for the various criteria has not been accomplished analytically. In general for models that are nonlinear in the parameters some numeric approximation to the integral is necessary (Chaloner and Verdinelli

1995).

Sándor and Wedel (2001) and Kessels, Goos and Vandebroek (2006) used a Monte Carlo estimate of the integral from $1,000$ random draws of the prior. Such estimates are known to converge to the true value of the integral at a rate proportional to the square root of the number of draws. This necessitates a large number of draws to reduce the sample-to-sample variability to the point where different samples do not lead to different design choices. This approach is costly because the computing time for the Bayesian design is then roughly $1,000$ times longer than the computing time for the local design, that is, the design for one prior parameter vector.

To solve integrals related to a multivariate normal distribution for the construction of choice designs, Sándor and Wedel (2002) utilized samples based on orthogonal arrays (Tang 1993) and Sándor and Wedel (2005) constructed quasi-Monte Carlo samples (Hickernell et al. 2000). In several cases, estimates using these methods are more efficient than Monte Carlo estimates so that it is possible to employ smaller samples to obtain the same accuracy (Sándor and András 2004; Sándor and Train 2004). There is also an extensive literature on quadrature, which is another approach to numerical integration. However, for integrals of functions in more than four dimensions, Monte Carlo estimates tend to outperform quadrature estimates (Geweke 1996; Monahan and Genz 1997).

A 20-point set

We propose to approximate the integrals in (2.2), (2.3), (2.4) and (2.6) with a designed sample of only 20 parameters. Assuming that the prior variance-covariance matrix $\boldsymbol{\Sigma}_0$ is the identity matrix, the multivariate normal distribution is spherically symmetric around the prior mean. As a result, every parameter has the same density on a k-dimensional hypersphere of a given radius. The 20 prior parameters are uniformally distributed on such a sphere. In this way, they sample the different directions away from the prior mean fairly.

For the comparison example, the designed sample of 20 parameters yields an approximation that is worse than the Monte Carlo sample of $1,000$ draws. However, in the computation of Bayesian optimal designs, it is not necessary for the approximation of the integral to be accurate. All that is required is that the sign of the difference from a rough approximation corresponding to two slightly different designs matches the sign of the difference from a better approximation. With the plot in Figure 2.3 we illustrate that the systematic 20-point sample and the Monte Carlo sample largely agree on design improvements in a random start.

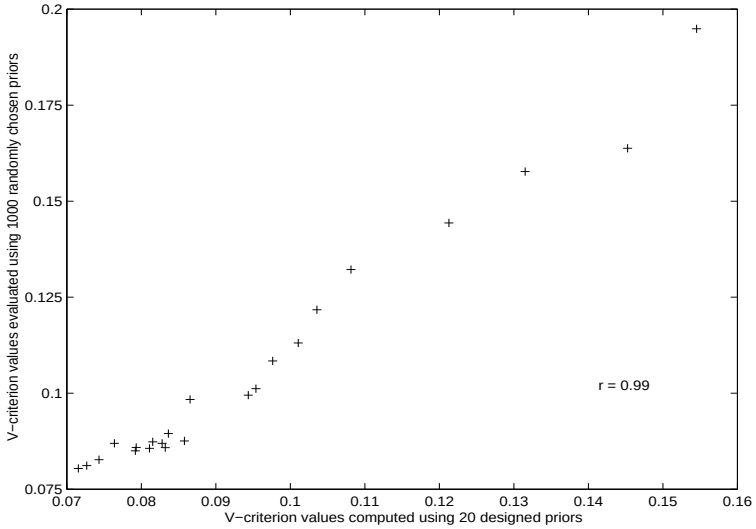


Figure 2.3: $\mathcal{V}_B$-criterion values according to the $1,000$-point Monte Carlo sample versus the systematic 20-point sample and correlation between them. The points represent the course of one try of the coordinate-exchange algorithm for the two-alternative designs using the 20-point sample.

The plot compares the $\mathcal{V}_B$-criterion value for the Monte Carlo sample with the $\mathcal{V}_B$-criterion value for the systematic 20-point sample. It depicts the course of one random start of the coordinate-exchange algorithm for the two-alternative designs. A random starting design is thereby monotonically improved by making a sequence of changes, each of which improves the $\mathcal{V}_B$-criterion value for the systematic 20-point sample. By re-evaluating each of these changes with the $\mathcal{V}_B$-criterion value for the Monte Carlo sample, we find out whether every change also leads to an improvement using the better

approximation.

The starting design is represented by the point at the top right of the plot, which of all points has the highest or worst $\mathcal{V}_B$-criterion value according to the 20-point sample as well as the Monte Carlo sample. After making one change in the original design, the second point from the top right shows an improvement in the $\mathcal{V}_B$-criterion value for both samples. The points proceed from the top right to the bottom left of the plot. The point at the bottom left denotes the final and best design produced in the random start. Note that this point has the lowest or best $\mathcal{V}_B$-criterion value as approximated by both samples. Also note that the drop in the $\mathcal{V}_B$-criterion value is not monotonic, indicating that the two approximations are not in complete agreement about the $\mathcal{V}_B$-criterion value of each change in the sequence.

Still, the agreement between the $\mathcal{V}_B$-criterion value for the Monte Carlo sample and the $\mathcal{V}_B$-criterion value for the systematic 20-point sample is clear from a correlation of 99%. Similar correlations are obtained using the coordinate-exchange algorithm with every other design criterion and for a larger choice set size. However, this does not imply that designs that are optimal using the systematic 20-point sample are also optimal with respect to the Monte Carlo sample. The plot in Figure 2.4 demonstrates this.

Like the plot in Figure 2.3, the plot in Figure 2.4 displays the $\mathcal{V}_B$-criterion value for the Monte Carlo sample versus the $\mathcal{V}_B$-criterion value for the systematic 20-point sample. Now each point in the plot represents the best two-alternative design found in a single random start of the coordinate-exchange algorithm. Again, the algorithm used the $\mathcal{V}_B$-criterion value for the 20-point sample to generate the designs and the $\mathcal{V}_B$-criterion value for the Monte Carlo sample to re-evaluate them. From the plot, we see that the worst design by both $\mathcal{V}_B$-criterion values is the same. On the other hand, the best design according to the $\mathcal{V}_B$-criterion value for the 20-point sample differs from the best design indicated by the $\mathcal{V}_B$-criterion value for the Monte Carlo sample.

In this case, the correlation between the $\mathcal{V}_B$-criterion values for the Monte Carlo sample and the $\mathcal{V}_B$-criterion values for the 20-point sample from the

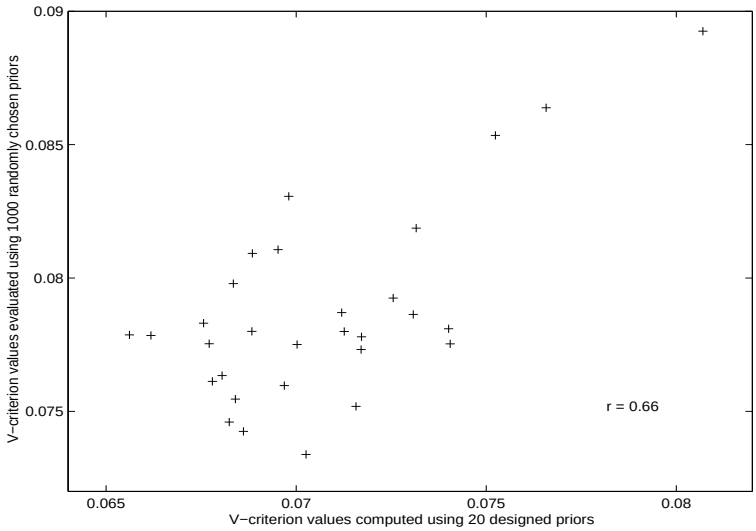


Figure 2.4: $\mathcal{V}_B$-criterion values according to the $1,000$-point Monte Carlo sample versus the systematic 20-point sample and correlation between them. The points correspond to two-alternative designs produced by different tries of the coordinate-exchange algorithm using the 20-point sample.

different tries is only 66%. This result also applies to the other design criteria and larger choice set sizes. The fact that the correlation is not close to 100% means that it is important to check each random start using the 20-point sample with one calculation of the objective function using the Monte Carlo sample. Therefore, our approach is an adaptive one in which we re-evaluate the Bayesian designs from the 20-point sample after each try using the Monte Carlo sample. The design with the best criterion value in terms of the $1,000$ draws is then selected.

Note that, if the correlation were near 100%, it would not be necessary to check the designs. On the other hand, if the correlation were not fairly large, then the adaptive approach would not work because designs using the 20-point sample would not be substantially better than random designs. Also, observe that thanks to the decrease in the number of prior parameters from $1,000$ to 20 during a try we save up to 98% of the computational work!

Constructing a small sample of prior parameters

For any choice design problem, we can construct a small set of prior parameters based on minimum potential designs or space filling designs created in JMP 6. The points of these designs are uniformally distributed on a k-dimensional hypersphere at a radius of one away from the zero vector. So on the sphere, the minimum distance to a neighboring point from any of the design points is roughly the same for all the points.

To understand how minimum potential designs are created, consider n points on a k-dimensional sphere around the zero vector. Each point, p, has levels between $[-1, 1]$ for k continuous factors and is denoted as $(z_{p1}, ..., z_{pk})$. Let d_{ef} be the distance between the eth and fth points. That is,

$$d_{ef} = \sqrt{\sum_{i=1}^{k} (z_{ei} - z_{fi})^2}. \tag{2.17}$$

The optimization problem is to find the $n \times k$ values of z_{pi} that minimize E_{pot}, the potential energy of the system:

$$E_{pot} = \sum_{e=1}^{n-1} \sum_{f=e+1}^{n} \left(d_{ef}^2 + \frac{1}{d_{ef}} \right). \tag{2.18}$$

Here, d_{ef}^2 is proportional to the energy stored in a spring when you pull it and $1/d_{ef}$ is the potential energy between two like charged particles. When the distance between two points increases, d_{ef}^2 increases. When the distance between two points decreases, $1/d_{ef}$ increases. To visualize this, Figure 2.5 shows a plane with 3 design points. Each point has springs attached to the other two points. The springs pull the points together. However, each point is also positively charged and the charges repel to push the points apart. The result is that the points end up forming an equilateral triangle.

For the comparison example, the minimum potential design with 20 points in a 5-dimensional space appears in Table 2.3. These points lie on a sphere of a radius of one around $[0, 0, 0, 0, 0]'$. The minimum distance for each point to the nearest point is 1.171. If this interpoint distance seems too large, then it can be reduced by increasing the number of points.

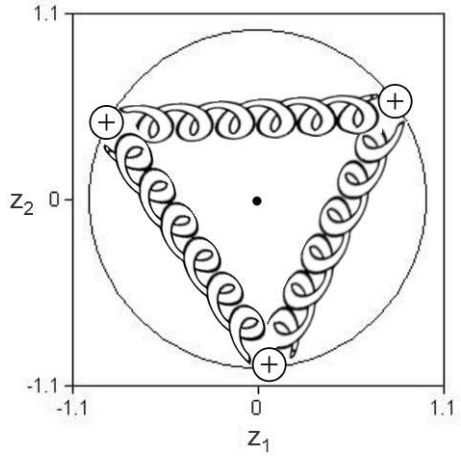


Figure 2.5: Three equally spaced points on the circumference of a circle.

Table 2.3: Minimum potential design of 20 points in 5 continuous factors for the comparison example.

Design point	z_1	z_2	z_3	z_4	z_5	Minimum distance	Nearest point	Radius
1	−0.17642	−0.57290	−0.19875	0.74536	−0.19600	1.17076	15	0.99281
2	−0.21775	0.81588	0.32619	−0.30104	−0.28759	1.17075	19	0.99281
3	−0.54891	−0.28739	−0.29445	0.17376	0.70655	1.17076	8	1.00000
4	−0.57116	0.06703	−0.27064	−0.77093	0.04122	1.17074	8	1.00000
5	−0.20011	−0.19572	−0.17339	−0.25973	−0.90384	1.17074	20	0.99281
6	0.00117	0.10528	0.59690	0.49371	−0.62360	1.17075	5	1.00000
7	−0.01228	0.13614	0.39319	−0.47950	0.76785	1.17076	13	0.99280
8	0.00528	−0.87552	−0.10638	−0.43810	0.15165	1.17074	4	0.99281
9	0.75353	−0.47946	−0.01214	0.10921	−0.43617	1.17076	16	1.00000
10	0.58274	0.19380	−0.32178	−0.71016	−0.08827	1.17075	20	0.99281
11	0.73699	0.47141	0.45742	0.07033	0.14296	1.17075	10	1.00000
12	−0.79511	−0.25333	0.54158	−0.02905	−0.04767	1.17077	13	0.99281
13	0.19427	−0.53359	0.65989	0.32602	0.36850	1.17075	17	1.00000
14	−0.00619	0.71761	−0.49688	−0.08192	0.48104	1.17075	2	1.00000
15	0.01039	−0.23327	−0.96643	−0.04302	−0.09815	1.17075	16	1.00000
16	0.60646	−0.18338	−0.34715	0.29484	0.61963	1.17075	15	0.99281
17	−0.19392	0.43870	0.30072	0.70409	0.42020	1.17075	13	0.99281
18	0.40102	0.49636	−0.41417	0.47335	−0.43591	1.17075	15	0.99281
19	−0.74200	0.35148	−0.35744	0.34145	−0.28555	1.17075	2	1.00000
20	0.17200	−0.17915	0.68369	−0.61868	−0.29685	1.17074	5	1.00000

To properly approximate the prior distribution with a 20-point sample from the points of a minimum potential design, it is necessary to rescale these points for the prior variance-covariance matrix and the prior mean. If there is no correlation between the prior coefficients or $\mathbf{\Sigma}_0 = \sigma_0^2 \mathbf{I}_k$, then the 20-point sample lies on a sphere with a radius that is proportional to the standard deviation σ_0. Now, the effectiveness of the 20-point sample in the adaptive

algorithm depends on the radius specified, or the number of standard deviations away from the prior mean. That is to say, a well-chosen radius requires fewer random starts to reach the global optimum. To find the best radius for a spherical 20-point sample for any choice design problem, one could proceed as follows:

1. Do a number of random starts of the adaptive algorithm for each of three radii,

2. Fit a quadratic function to the minimum criterion value found at each radius,

3. Choose the radius that is the minimum of the quadratic function.

For the comparison example, we performed 10 random starts for a radius of 1, 2 and 3. Recall that $\sigma_0 = 1$ for this example. The result for the $\mathcal{V}_B$-optimality criterion connected with two-alternative designs appears in Figure 2.6. Fitting a quadratic model to the minima results in a radius slightly larger than 2. We chose however a radius of 2 for simplicity. To illustrate the value of selecting a good radius, we compared the estimated expected efficiencies per number of tries of the two-alternative $\mathcal{V}_B$-optimal designs using the 20-point samples for the radii 1 and 2, respectively. The plots based on 250 tries appear in Figure 2.7. We clearly observe the higher expected efficiencies in case a radius of 2 is utilized as opposed to a radius of 1. We obtained similar results for any other optimality criterion in combination with any choice set size.

However, computing the "best" radius is not absolutely necessary. The heuristic of choosing a sphere radius that is twice the prior standard deviation worked well in all the examples we tried. The critical part of the adaptive algorithm is that for each random start using the 20-point sample, one checks the resulting design with the larger Monte Carlo sample. So, no matter what radius one chooses, one will have a monotonically improving set of designs as the number of random starts increases. Still, choosing a good radius increases the speed of the improvement over the random starts.

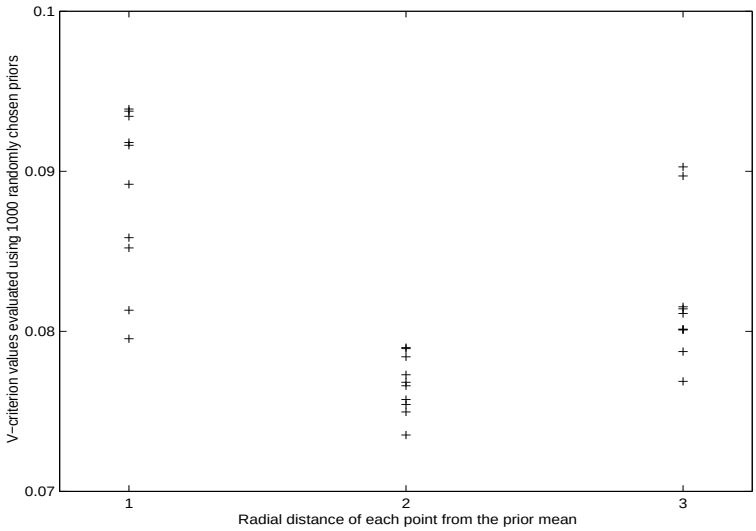


Figure 2.6: $\mathcal{V}_B$-criterion values of two-alternative designs from 10 random starts of the adaptive algorithm using the 20-point samples for the radii 1, 2 and 3.

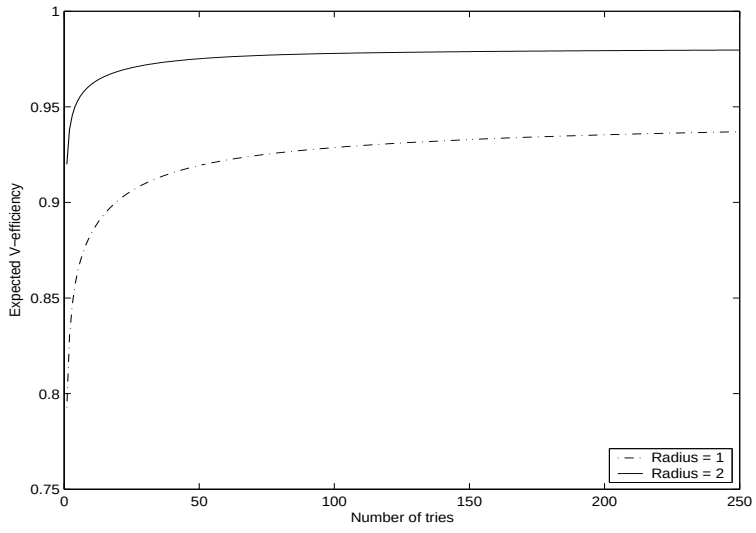


Figure 2.7: Estimated expected efficiencies per number of tries of the two-alternative $\mathcal{V}_B$-optimal designs computed using the adaptive algorithm with the 20-point samples for the radii 1 and 2.

2.5 Computation of large choice designs

The speed of the adaptive algorithm makes the computation of Bayesian optimal designs feasible for more challenging problems of larger dimensions than the rather small comparison example $3^2 \times 2/24$. We illustrate this with the construction of designs of two classes: $5\times3\times2^3/2/15$ and $5\times3\times2^3/3/10$, jointly referred to as $5\times3\times2^3/30$. The designs consist of 30 profiles, grouped in 15 choice sets of size two for the first class and 10 choice sets of size three for the second class. The profiles are configured from five attributes, one of which has five levels, another of which has three levels and the three others of which have two levels. So in total, there are $5 \times 3 \times 2^3 = 120$ candidate profiles. This candidate set is much larger than the candidate set of 18 profiles employed in the comparison example.

For the two classes, we constructed designs using the $\mathcal{D}_B$- and $\mathcal{V}_B$-optimality criteria. The $\mathcal{D}_B$-optimality criterion is the most popular criterion of the estimation-based design criteria. For the prediction-based design criteria, we prefer the $\mathcal{V}_B$-optimality criterion since it seeks to minimize the average prediction variance over the design region χ and, as we showed in Section 2.4.2, its criterion value can be computed more efficiently than the $\mathcal{G}_B$-criterion value. For the two-alternative design class, χ consists of $Q = \binom{120}{2} = 7,140$ choice sets, or $14,280$ profiles, whereas for the three-alternative design class, it comprises $Q = \binom{120}{3} = 280,840$ choice sets, or $842,520$ profiles. Compare these numbers with the 306 profiles and $2,448$ profiles for the two- and three-alternative designs of the comparison example.

The number of parameter values, k, using effects-type coding is nine. As prior parameter distribution, we implemented the multivariate normal distribution $\pi(\boldsymbol{\beta}) = \mathcal{N}(\boldsymbol{\beta}|\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$, with $\boldsymbol{\beta}_0 = [-1, -0.5, 0, 0.5, -1, 0, -1, -1, -1]'$ and $\boldsymbol{\Sigma}_0 = \mathbf{I}_9$. To obtain the designs for the $\mathcal{D}_B$- and $\mathcal{V}_B$-optimality criteria, we performed $1,000$ tries of the adaptive algorithm for each criterion. We therefore utilized a constructed 20-point sample for the design generation and a random $1,000$-point sample for the design evaluation. Again, we carried out all computations in MATLAB 7 by means of a Dell personal computer with a 1.60 GHz Intel Processor and 2 GB RAM.

The $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs for the $5 \times 3 \times 2^3/30$ example appear in Tables A4 and A5 of the Appendix. Their criterion values appear in Table 2.4. For both optimality criteria, we notice a decrease or an improvement in the values as the choice set size goes from two to three. The result that the performance in terms of prediction improves with the choice set size was also noted by Sándor and Wedel (2002) and Kessels, Goos and Vandebroek (2006). The $\mathcal{V}_B$-criterion values for the comparison example in Table 2.1 further confirm this. However, we remain undecided as to the efficiency of the $\mathcal{D}_B$-optimal designs with respect to the choice set size. In contrast to Table 2.4 where the $\mathcal{D}_B$-criterion values decrease with larger choice sets, the $\mathcal{D}_B$-criterion values in Table 2.1 for the comparison example increase with larger choice sets.

Table 2.4: $\mathcal{D}_B$- and $\mathcal{V}_B$-criterion values of the two- and three-alternative $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs for the $5 \times 3 \times 2^3/30$ design example.

Optimal design	# Alternatives	
	2	3
$\mathcal{D}_B$	1.18591	1.13639
$\mathcal{V}_B$	0.23901	0.21065

The computing times for one try of the adaptive algorithm to generate the two- and three-alternative $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs for the $5 \times 3 \times 2^3/30$ example appear in Table 2.5. The huge design region for the three-alternative designs results in a running time of several minutes per try for the $\mathcal{V}_B$-optimality criterion. The computation of the $\mathcal{V}_B$-optimal designs for this large example would have taken months using MCMF.

Table 2.5: Computing times for one try of the adaptive algorithm to generate the two- and three-alternative $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs for the $5 \times 3 \times 2^3/30$ design example. The times are expressed in hours:minutes:seconds.

Design criterion	# Alternatives	
	2	3
$\mathcal{D}_B$	00:00:08	00:00:14
$\mathcal{V}_B$	00:00:15	00:04:05

2.6 Conclusion

In this chapter, we propose an adaptive algorithm for producing $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal choice designs as an alternative to the Monte Carlo modified Fedorov algorithm (MCMF) employed by Kessels, Goos and Vandebroek (2006). Kessels, Goos and Vandebroek (2006) had shown that $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs outperform $\mathcal{D}_B$- and $\mathcal{A}_B$-optimal designs for response prediction, which is central in choice experiments. However, using MCMF computing $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs is even more cumbersome than searching for $\mathcal{D}_B$- and $\mathcal{A}_B$-optimal designs so that they suggested implementing the $\mathcal{D}_B$-optimality criterion in practice.

Unlike MCMF, the new adaptive algorithm makes the construction of $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs practical and it allows the $\mathcal{D}_B$-, $\mathcal{A}_B$-, $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs to be embedded in web-based conjoint choice studies with individualized designs for the respondents. We prefer using $\mathcal{V}_B$-optimal designs since they minimize the average prediction variance and can be computed faster than $\mathcal{G}_B$-optimal designs. In general, the main improvement of the adaptive algorithm over MCMF is the approximation of the normal prior distribution by a designed sample of 20 parameter vectors instead of a Monte Carlo sample of $1,000$ random draws. This saves up to 98% of the computational work within each try of the algorithm. Nevertheless, we re-evaluate the designs produced by each try using the Monte Carlo sample and adapt the design selection accordingly. This led us to call our method the adaptive algorithm.

To further speed up the design generation, the adaptive algorithm also uses a coordinate-exchange algorithm rather than a modified Fedorov algorithm. A coordinate-exchange approach saves time by avoiding the creation and use of a candidate set that grows exponentially with the number of attributes and attribute levels studied. Thus, the time savings of the coordinate-exchange algorithm increase with the number of profiles, attributes and attribute levels. As a last way to accelerate the computations for any optimality criterion, the adaptive algorithm incorporates an update formula to economically calculate the optimality criterion values of designs.

The computational speed of the adaptive algorithm makes the use of individualized Bayesian optimal designs in web-based surveys possible. To examine what is the best way to do this, is beyond the scope of this chapter. We expect, however, that such an approach would allow an efficient estimation of mixed logit (Sándor and Wedel 2002) and latent class models (Andrews, Ainslie and Currim 2002; Train 2003) that aim at modelling consumer heterogeneity. Another topic for further research is the construction of designs for choice experiments in which one suspects correlation between parameter coefficients. In that case, the multivariate normal prior distribution is elliptically symmetric around the prior mean. The small designed sample of parameters from a minimum potential design should then be rescaled to lie on a k-dimensional ellipsoid. Lastly, the efficiency of optimal designs with respect to the choice set size might be further investigated.

Appendix. Choice design tables

Table A1: Two-alternative Bayesian optimal designs for the $3^2 \times 2/24$ example.

		$\mathcal{D}_B$			$\mathcal{A}_B$			$\mathcal{G}_B$			$\mathcal{V}_B$		
Choice	Alt	Attr			Attr			Attr			Attr		
set		1	2	3	1	2	3	1	2	3	1	2	3
1	I	1	2	2	2	1	2	1	2	2	1	3	2
	II	2	1	1	1	2	1	3	1	2	2	1	2
2	I	3	1	1	1	2	2	1	2	1	1	2	2
	II	2	3	2	2	1	1	2	3	1	2	3	2
3	I	2	3	2	2	2	2	2	1	1	2	2	2
	II	1	2	2	1	1	1	3	2	2	1	1	2
4	I	3	3	2	2	2	1	2	2	1	2	2	1
	II	2	2	1	3	1	1	1	3	1	1	3	2
5	I	2	2	2	2	1	1	1	2	1	1	1	1
	II	1	3	2	1	2	1	2	1	2	2	2	2
6	I	1	2	2	2	3	2	3	3	1	2	1	1
	II	3	1	2	3	2	1	2	1	1	3	1	2
7	I	3	1	2	1	2	2	2	3	2	1	2	2
	II	1	3	1	2	3	2	3	1	1	3	1	2
8	I	1	2	1	1	3	1	3	2	1	1	1	1
	II	2	1	2	3	1	2	1	1	1	3	3	2
9	I	2	2	2	1	1	1	3	1	2	1	2	1
	II	1	1	1	3	1	1	1	3	1	2	1	2
10	I	2	2	1	3	3	2	2	2	1	2	3	2
	II	1	1	1	2	1	1	3	3	2	3	1	1
11	I	3	2	1	2	2	2	2	2	2	3	2	2
	II	2	1	1	1	3	2	1	1	1	2	1	2
12	I	2	1	1	2	2	1	1	1	2	1	2	2
	II	3	2	2	1	1	1	2	1	1	1	3	1

Table A2: Three-alternative Bayesian optimal designs for the $3^2 \times 2/24$ example.

		$\mathcal{D}_B$			$\mathcal{A}_B$			$\mathcal{G}_B$			$\mathcal{V}_B$		
Choice	Alt	Attr			Attr			Attr			Attr		
set		1	2	3	1	2	3	1	2	3	1	2	3
1	I	2	1	1	1	2	1	2	1	1	3	1	1
	II	1	3	1	2	1	2	1	2	2	1	3	2
	III	1	2	2	1	3	2	1	3	1	2	2	1
2	I	2	3	2	1	1	1	2	2	1	3	1	1
	II	1	2	1	1	2	2	1	3	2	2	2	2
	III	3	1	2	2	2	1	3	2	2	3	2	2
3	I	1	2	1	2	2	2	2	3	1	3	1	2
	II	2	3	1	1	2	2	2	1	2	2	3	2
	III	1	3	2	1	1	1	3	3	1	2	2	1
4	I	3	2	2	1	2	1	1	1	1	1	3	1
	II	2	1	1	1	1	1	3	1	2	2	1	1
	III	1	3	2	2	1	1	2	3	2	1	2	2
5	I	2	1	2	2	2	1	3	2	1	1	1	1
	II	2	2	2	3	1	1	3	3	1	2	3	2
	III	1	1	1	1	3	1	1	3	1	3	2	2
6	I	3	3	2	1	2	1	1	2	1	3	3	2
	II	2	2	1	2	3	1	2	3	2	1	3	1
	III	1	1	1	3	3	2	2	3	1	2	1	2
7	I	3	1	2	2	3	2	1	1	1	1	3	2
	II	2	1	1	3	1	1	1	2	2	2	3	2
	III	3	2	1	3	3	1	2	1	2	1	2	2
8	I	1	2	1	2	1	1	2	1	1	3	2	1
	II	3	1	1	3	2	1	1	1	2	3	1	2
	III	2	2	2	1	2	1	2	2	1	1	1	1

Table A3: Four-alternative Bayesian optimal designs for the $3^2 \times 2/24$ example.

		$\mathcal{D}_B$			$\mathcal{A}_B$			$\mathcal{G}_B$			$\mathcal{V}_B$		
Choice	Alt	Attr			Attr			Attr			Attr		
set		1	2	3	1	2	3	1	2	3	1	2	3
1	I	2	2	2	2	3	2	2	3	1	3	1	1
	II	2	1	1	3	2	2	2	2	1	1	2	1
	III	3	1	2	2	2	2	1	3	1	2	1	1
	IV	3	2	1	1	1	1	1	2	2	3	2	2
2	I	2	1	2	1	3	1	3	2	1	3	2	2
	II	3	1	1	2	3	2	2	1	2	1	1	1
	III	3	2	2	1	3	2	1	1	1	2	3	2
	IV	1	2	1	2	1	1	3	1	1	2	2	2
3	I	2	2	1	3	1	1	1	3	2	2	3	1
	II	1	1	1	3	2	1	1	1	2	1	3	2
	III	1	3	2	1	2	2	1	2	1	2	1	2
	IV	3	1	2	2	1	1	2	2	2	1	2	1
4	I	1	2	1	2	2	1	2	1	1	3	3	1
	II	2	1	1	1	1	1	3	1	2	3	1	1
	III	1	3	1	3	1	2	1	2	2	1	3	1
	IV	2	3	2	1	3	2	2	3	2	2	2	2
5	I	3	2	2	3	2	2	2	3	1	2	2	1
	II	2	2	2	1	2	1	3	3	1	2	1	1
	III	2	3	2	3	1	1	3	1	1	1	3	2
	IV	1	1	1	1	1	1	3	3	2	3	1	2
6	I	1	2	2	1	3	1	3	1	2	2	3	2
	II	2	3	1	2	1	2	1	2	2	2	2	1
	III	1	3	1	2	3	1	2	1	2	1	2	2
	IV	2	1	1	1	2	1	1	1	1	1	3	1

Table A4: Two-alternative $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs for the $5 \times 3 \times 2^3/30$ example.

		$\mathcal{D}_B$					$\mathcal{V}_B$				
Choice	Alt	Attr					Attr				
set		1	2	3	4	5	1	2	3	4	5
1	I	5	2	1	2	2	4	1	1	2	2
	II	4	3	2	1	2	5	1	2	2	1
2	I	3	2	2	1	1	2	1	1	1	2
	II	2	3	1	2	1	4	2	2	1	2
3	I	5	1	2	1	1	3	3	1	2	1
	II	1	2	1	2	1	5	2	1	2	1
4	I	4	1	1	1	1	4	3	1	1	2
	II	2	2	1	1	1	3	2	1	1	2
5	I	3	1	1	2	1	5	3	2	2	1
	II	4	2	2	1	2	2	2	2	1	1
6	I	1	2	2	1	1	2	1	2	2	1
	II	2	1	2	1	1	3	2	2	2	2
7	I	2	2	2	2	1	3	1	2	2	1
	II	3	1	1	2	2	2	3	2	1	2
8	I	1	1	1	1	1	1	3	1	2	2
	II	4	3	1	2	2	3	3	1	2	2
9	I	2	1	1	1	2	2	3	2	2	1
	II	5	2	2	2	1	1	1	2	1	1
10	I	5	3	1	1	2	4	3	2	2	2
	II	4	2	2	1	1	3	3	2	1	2
11	I	1	1	2	2	2	2	2	2	1	1
	II	3	2	2	2	2	5	1	1	1	2
12	I	3	1	2	1	1	1	2	2	1	1
	II	5	1	2	1	2	2	2	2	1	2
13	I	1	3	2	1	1	4	1	2	2	1
	II	3	2	1	2	1	1	2	1	2	1
14	I	4	1	2	2	1	4	1	2	1	1
	II	3	3	2	2	2	1	3	2	2	1
15	I	3	3	2	2	1	5	3	2	1	2
	II	4	2	1	1	2	4	2	1	2	1

Table A5: Three-alternative $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs for the $5 \times 3 \times 2^3/30$ example.

		$\mathcal{D}_B$					$\mathcal{V}_B$				
Choice	Alt	Attr					Attr				
set		1	2	3	4	5	1	2	3	4	5
1	I	4	3	1	2	2	5	1	2	1	1
	II	1	1	1	1	1	4	3	2	1	2
	III	2	2	2	1	1	3	2	2	1	2
2	I	2	3	2	1	2	1	1	1	2	1
	II	1	3	1	2	1	5	2	2	1	2
	III	5	2	2	1	2	3	3	2	1	2
3	I	4	3	2	1	1	1	2	2	1	2
	II	3	3	1	1	2	3	2	2	2	1
	III	2	2	1	1	1	4	1	2	2	1
4	I	2	3	1	1	1	4	2	1	1	1
	II	1	2	1	1	1	2	3	1	2	1
	III	5	3	1	2	2	1	1	1	1	1
5	I	3	1	2	2	1	5	3	2	1	2
	II	1	2	2	2	1	1	3	1	1	1
	III	2	2	2	2	2	4	2	1	2	1
6	I	2	1	2	2	1	3	3	2	1	1
	II	3	2	2	1	2	2	2	1	1	2
	III	4	1	2	1	2	5	3	1	1	2
7	I	4	2	1	2	2	2	2	1	1	2
	II	3	1	2	1	1	5	3	2	2	1
	III	5	3	2	2	1	4	2	2	1	2
8	I	4	1	1	2	1	1	2	1	2	1
	II	2	3	1	1	2	3	1	1	1	2
	III	3	3	2	2	1	2	1	2	1	1
9	I	4	2	2	2	1	5	1	2	2	1
	II	1	3	2	2	1	2	3	2	1	1
	III	5	1	1	2	2	3	3	1	2	2
10	I	3	2	1	2	1	1	3	2	1	2
	II	5	1	2	1	1	2	3	1	2	2
	III	1	3	1	1	2	3	1	1	1	2

Chapter 3

Recommendations on the use of Bayesian choice designs

This chapter has been submitted as

☞ Kessels, R., Jones, B., Goos, P. and Vandebroek M. (2006). Recommendations on the use of Bayesian optimal designs for choice experiments.

Abstract

In this chapter, we argue that some of the prior parameter distributions used in the literature for the construction of Bayesian optimal designs are internally inconsistent. We rectify this error and provide practical advice on how to properly specify the prior parameter distribution. Also, we present two pertinent examples to illustrate that Bayesian optimal designs generally outperform utility-neutral optimal designs that are based on linear design principles.

3.1 Introduction

Choice experiments have become an increasingly popular method to understand consumers' preference structures for the attributes of a product or service. In such experiments respondents make a sequence of choices. In each case they indicate their preferred product or service among a choice set of alternatives or profiles. A profile is thereby characterized by a combination of attribute levels. The design of a choice experiment comprises a select number of choice sets administered to each respondent. The aim of a choice experiment is to estimate the importance of each attribute and their levels based on the respondents' preferences. The estimates are then used to mimic real marketplace choices by making predictions about consumers' future purchases.

The question of how to design efficient choice experiments has received a great deal of attention recently. Designing an efficient choice experiment involves selecting those choice sets that result in a precisely estimated model providing accurate predictions. At present, two design approaches are prevalent: the Bayesian design approach and the linear design approach. We review these current practices for setting up choice experiments.

Bayesian choice designs have so far been constructed for the multinomial logit model (McFadden 1974). This discrete choice model predicts for profile j, $j = 1, ..., J$, in choice set s, $s = 1, ..., S$, the probability that people prefer it: $p_{js} = e^{\mathbf{x}'_{js}\boldsymbol{\beta}} / \sum_{t=1}^{J} e^{\mathbf{x}'_{ts}\boldsymbol{\beta}}$. Here, $\mathbf{x}_{js}$ is a $k \times 1$ vector of the attribute levels of profile j in choice set s and $\boldsymbol{\beta}$ is a $k \times 1$ vector of parameter values. The multinomial logit probability is derived from people's latent utility for profile j in choice set s: $U_{js} = \mathbf{x}'_{js}\boldsymbol{\beta} + \varepsilon_{js}$ where ε_{js} is an i.i.d. extreme value error term. Since the multinomial logit model is nonlinear in the parameters, like all other choice models, the quality of a given design depends on the unknown parameter vector. The Bayesian design approach deals with this problem by assuming a prior distribution of likely parameters. It thereby takes into account the uncertainty on the proposed parameters. To date, most of the Bayesian research focus has been on designs for main-effects models.

Sándor and Wedel (2001) were the first to introduce the Bayesian design procedure in the choice design literature. They generated Bayesian designs using the $\mathcal{D}$-optimality criterion for the multinomial logit model. This design criterion seeks to minimize the determinant of the variance-covariance matrix of the parameter estimators. In the Bayesian framework, it is referred to as the $\mathcal{D}_B$-optimality criterion. Sándor and Wedel (2001) showed that the $\mathcal{D}_B$-optimal designs generally outperform the locally $\mathcal{D}_P$-optimal designs which are based on a point estimate for the unknown parameter vector (Huber and Zwerina 1996). Sándor and Wedel (2005) continued the Bayesian approach to construct so-called heterogeneous $\mathcal{D}_B$-optimal designs that include several different designs that are each offered to different respondents.

Kessels, Goos and Vandebroek (2006) expanded the work on Bayesian choice designs by also considering other design criteria than the commonly used $\mathcal{D}_B$-optimality criterion. They compared the $\mathcal{D}_B$- and $\mathcal{A}_B$-optimality criteria with the $\mathcal{G}_B$- and $\mathcal{V}_B$-optimality criteria for the multinomial logit model. The $\mathcal{D}_B$- and $\mathcal{A}_B$-optimality criteria concentrate on producing precise estimates, whereas the $\mathcal{G}_B$- and $\mathcal{V}_B$-optimality criteria focus on providing precise predictions, which is key in choice experiments. Using a simulation study, Kessels, Goos and Vandebroek (2006) demonstrated that the $\mathcal{D}_B$- and $\mathcal{A}_B$-optimal designs actually produce more precise estimates and that the $\mathcal{G}_B$- and $\mathcal{V}_B$-optimal designs produce better predictions. Also, they showed that the $\mathcal{D}_B$-optimal designs perform reasonably well in terms of prediction.

To quickly generate the Bayesian designs, Kessels et al. (2006) developed an adaptive algorithm. The high speed of this algorithm stems from the use of a small designed sample of prior parameters to approximate the prior distribution, Meyer and Nachtsheim's (1995) coordinate-exchange algorithm, and an update approach to economically calculate the criterion values of designs that differ only in one profile from another design. Kessels et al. (2006) recommended using $\mathcal{V}_B$-optimal designs primarily because they are faster to compute. Also, Kessels et al. (2006) preferred minimizing the average prediction variance to minimizing the maximum prediction variance over the design region, as the $\mathcal{V}_B$- and $\mathcal{G}_B$-optimal designs do, respectively.

Currently, however, linear design principles are still used to construct designs for choice experiments. Such designs are based on an implicit assumption that the respondents are indifferent to all attribute levels, and thus to all alternatives. Moreover, there is no uncertainty associated with the indifference. This is equivalent to adopting a zero prior parameter vector with zero prior variance for the multinomial logit model. The designs are therefore referred to as utility-neutral designs and they are utility balanced by assumption (Huber and Zwerina 1996). Utility-neutral designs for main effects as well as main effects plus interactions have been discussed at length. To generate them, Kuhfeld and Tobias (2005) proposed a $\mathcal{D}$-efficient factorial design algorithm implemented in the SAS `%MktEx` macro. This algorithm combines Cook and Nachtsheim's (1980) modification of Fedorov's (1972) exchange algorithm, the coordinate-exchange algorithm with simulated annealing, and a very large catalog of orthogonal arrays.

Street, Bunch and Moore (2001) and Street and Burgess (2004) followed a more theoretical approach providing generators to construct utility-neutral paired comparison designs for two-level attributes. In paired comparison designs profiles are arranged in choice sets of size two. The authors used the nonlinear Bradley-Terry model, that is the logit model for paired evaluations, for which they assumed zero prior parameter values. In most of this work, the focus was on the $\mathcal{D}$-optimality criterion, but Street, Bunch and Moore (2001) also computed $\mathcal{A}$-optimal pairs, which minimize the sum or the average of the variances of the parameter estimators. Furthermore, Burgess and Street (2003) derived $\mathcal{D}$-optimal utility-neutral designs for two-level attributes of any choice set size. Even more flexible designs allowing for attributes with any number of levels are elaborated in Burgess and Street (2005). Finally, Street, Burgess and Louviere (2005) showed that the theoretical strategies proposed in their aforementioned papers produce utility-neutral designs that are better than those based on common strategies.

The assumption of complete indifference among all alternatives underlying the utility-neutral designs is surely unrealistic. The Bayesian design approach is more practical since it incorporates all available prior information in the designs. Moreover, Bayesian optimal designs are on average more efficient than

utility-neutral optimal designs. In this chapter we show that this is indeed true using two design examples. Before doing so, we first rectify a misunderstanding with respect to the specification of the prior parameter distribution.

In a number of Bayesian design examples studied by Sándor and Wedel (2001), Kessels, Goos and Vandebroek (2006) and Kessels et al. (2006) where prior information on the parameter vector $\boldsymbol{\beta}$ from previous experiments is lacking, the specification of the prior distribution is impractical. In these examples Bayesian optimal designs are constructed for a multivariate normal distribution $\pi(\boldsymbol{\beta}) = \mathcal{N}(\boldsymbol{\beta}|\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$, where the elements in the prior mean $\boldsymbol{\beta}_0$ are equally spaced between -1 and 1 for each attribute and the prior variance-covariance matrix $\boldsymbol{\Sigma}_0$ is the identity matrix. In the next section, we show that these specifications of $\boldsymbol{\beta}_0$ and $\boldsymbol{\Sigma}_0$ conflict. Also, we provide some general recommendations on how to properly specify the prior parameter distribution $\pi(\boldsymbol{\beta}) = \mathcal{N}(\boldsymbol{\beta}|\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$ for any design case.

3.2 Guidance on correctly specifying the prior parameter distribution

We illustrate for the $3^2 \times 2/24$ design example, initiated in Kessels, Goos and Vandebroek (2006) and extended in Kessels et al. (2006), that the prior parameter distribution used to construct the Bayesian designs is unrealistic. In this example, the profiles are composed of two attributes at three levels and one attribute at two levels. The total number of design profiles is 24 and they have been arranged in choice sets of size two, three and four. Through effects-type coding, the number of elements, k, in the parameter vector $\boldsymbol{\beta}$ is 5. The prior parameter distribution exploited was the multivariate normal distribution $\pi(\boldsymbol{\beta}) = \mathcal{N}(\boldsymbol{\beta}|\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$. The parameter values of the prior mean $\boldsymbol{\beta}_0$ were evenly spaced between -1 and 1 for each attribute so that $\boldsymbol{\beta}_0 = [-1, 0, -1, 0, -1]'$, and the prior variance-covariance matrix $\boldsymbol{\Sigma}_0$ was the identity matrix $\mathbf{I}_5$. We explain why these specifications of $\boldsymbol{\beta}_0$ and $\boldsymbol{\Sigma}_0$ are contradictory.

First, however, a note should be made about the effects-type coding we adopt in this chapter. For a two-level attribute, Sándor and Wedel (2001), Kessels, Goos and Vandebroek (2006) and Kessels et al. (2006) coded the first level as -1 and the second level as 1 while specifying a prior mean parameter value of -1. In this way, a utility of 1 is attached to the first level and a utility of -1 to the second level so that the utilities decrease with the attribute levels. On the other hand, for attributes with more than two levels, the authors coded the levels such that the utilities increase with the levels given prior mean values that are equally spaced between -1 and 1. For example, for an attribute with three levels, the first level is coded as [1 0], the second level as [0 1] and the third level as [-1 -1]. Given prior mean values of $[-1,\ 0]'$ the utilities associated with the three levels are -1, 0 and 1, respectively. In order to have the utilities increase with the levels for all attributes, we change the coding for a two-level attribute to 1 for the first level and to -1 for the second level.

Consider now the two-alternative choice set in Table 3.1 for the $3^2 \times 2/24$ design example. This choice set is special in the sense that Alternative I consists of the worst possible levels for each attribute given the prior mean $\boldsymbol{\beta}_0 = [-1, 0, -1, 0, -1]'$ and Alternative II consists of the best possible attribute levels. As a result, Alternative II dominates the choice set. This can also be seen from the logit probabilities given $\boldsymbol{\beta}_0$. The probability that Alternative I is chosen is 0.00247 and the probability that Alternative II is chosen is 0.99753. These probabilities are most extreme meaning that there is no other two-alternative choice set for the $3^2 \times 2/24$ example with more extreme logit probabilities given $\boldsymbol{\beta}_0$. So, these probabilities imply very strong prior information. In other words, the prior mean is very informative about the overall attractiveness of the two alternatives in the choice set of Table 3.1.

Now, given the prior variance $\boldsymbol{\Sigma}_0 = \mathbf{I}_5$, the parameters $\boldsymbol{\beta}^1 = [0, 0, 0, 0, 0]'$ and $\boldsymbol{\beta}^2 = [-2, 0, -2, 0, -2]'$ are equally likely under the prior mean $\boldsymbol{\beta}_0$ and neither is improbable when drawn from a Monte Carlo sample. Using $\boldsymbol{\beta}^1$ the probabilities of choosing Alternatives I and II are 0.5 each. Using $\boldsymbol{\beta}^2$ the probabilities of choosing Alternatives I and II are 0.00001 and 0.99999, respectively. The differences in probabilities from using $\boldsymbol{\beta}^1$ and $\boldsymbol{\beta}^2$ seem to

Table 3.1: A two-alternative choice set for the $3^2 \times 2/24$ design example.

Alt	Attr		
	1	2	3
I	1	1	1
II	3	3	2

imply that not much prior information is assumed when $\boldsymbol{\Sigma}_0 = \mathbf{I}_5$ is used as the variance-covariance matrix of the prior distribution.

To illustrate a more extreme case where prior information is completely lacking, we ponder the parameters $\boldsymbol{\beta}^3 = [1, 0, 1, 0, 1]'$ and $\boldsymbol{\beta}^4 = [-3, 0, -3, 0, -3]'$. Similar to $\boldsymbol{\beta}^1$ and $\boldsymbol{\beta}^2$, these parameters are equally likely when $\boldsymbol{\beta}_0$ and $\mathbf{I}_5$ are used as prior mean and variance, and neither is improbable in a Monte Carlo sample. However, since they are further away from the prior mean, they are less plausible than $\boldsymbol{\beta}^1$ and $\boldsymbol{\beta}^2$. Using $\boldsymbol{\beta}^3$ the probabilities of choosing Alternatives I and II are 0.99753 and 0.00247, respectively. Using $\boldsymbol{\beta}^4$ the probabilities are reversed, essentially equaling 0 and 1.

Based on the above observations, the prior mean indicates that one has a substantial amount of prior knowledge about people's preferences for the alternatives in the choice set of Table 3.1. In fact, one has so much prior information that the choice set should not be included in the design. This is indeed the case when examining the optimal designs with choice sets of size two generated by Kessels et al. (2006). On the other hand, the prior variance implies that one has very little prior knowledge because the range of expected probabilities for the two alternatives in the choice set essentially goes from zero to one.

Hence, the prior parameter distribution $\pi(\boldsymbol{\beta}) = \mathcal{N}(\boldsymbol{\beta}|\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$, with $\boldsymbol{\beta}_0 = [-1, 0, -1, 0, -1]'$ and $\boldsymbol{\Sigma}_0 = \mathbf{I}_5$ is internally inconsistent. If one knows as much as the mean implies, then the variance should be smaller. If one knows as little as the variance implies, then the mean should be closer to the zero vector. Consequently, to specify a proper prior parameter distribution, one

has to choose between

1. an informative mean with a small variance and
2. a less informative mean with a larger variance.

The first option makes sense if you are augmenting a previous study, or verifying its results. In that case, the posterior mean and variance from the previous work can be used as the prior mean and variance for the new study.

The second strategy is appropriate in a case where no previous work has been done. Most often, one has some prior beliefs about the relative preferences for the attributes and its levels. It is sensible to incorporate these notions in the prior mean. However, if one is completely without intuition about what choices will be made by the market segment one is targeting, then the zero vector should be used as prior mean.

When dealing with ordinal attributes like for example the price of an apartment, the speed of a computer, the size of a house, and so forth, one has generally a clear idea about the overall predilection for the attribute levels. Utilities usually either increase or decrease when going from the low to the high setting of an ordinal attribute. It is wise to reflect this information in the prior mean.

To ensure that one uses an appropriate prior mean $\boldsymbol{\beta}_0$ for the prior parameter distribution $\pi(\boldsymbol{\beta}) = \mathcal{N}(\boldsymbol{\beta}|\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$ when no previous studies have been performed, we propose the following sanity check for $\boldsymbol{\beta}_0$:

1. List all possible choice sets of size two and compute the multinomial logit probabilities for each profile in these choice sets given $\boldsymbol{\beta}_0$.
2. Check whether the probabilities for all alternatives are reasonable. Do they match one's subjective probabilities or beliefs? Does one feel as confident about the alternatives as the logit probabilities imply?
3. (a) If yes, then $\boldsymbol{\beta}_0$ is a good choice.
   (b) If no, then choose a new prior mean in accordance with your understanding:

i. If the probabilities of the alternatives tend to overestimate one's beliefs, or one knows less than the probabilities indicate, then $\boldsymbol{\beta}_0$ should be taken closer to the zero vector. This draws the probabilities nearer to each other.

ii. If the probabilities underestimate one's understanding, or one knows more than the probabilities reveal, then $\boldsymbol{\beta}_0$ should be taken somewhat further from the zero vector. This pulls the probabilities more apart.

Note that overstating one's beliefs occurs more frequently than understating one's beliefs. Subsequently, verify whether the new prior mean is suitable by repeating the procedure.

Instead of going through all possible choice sets of size two, a more instant check on the suitability of the prior mean $\boldsymbol{\beta}_0$ is to examine only the choice set with the least attractive alternative and the most attractive alternative given $\boldsymbol{\beta}_0$. Assuming main-effects models, the least attractive alternative is composed by selecting the worst possible level for each attribute and the most attractive alternative is composed by selecting the best possible level for each attribute. As already mentioned, the choice set of Table 3.1 groups these alternatives for the $3^2 \times 2/24$ design example given the prior mean $\boldsymbol{\beta}_0 = [-1, 0, -1, 0, -1]'$. Once the choice set with the most extreme alternatives is constructed, the logit probabilities should be computed and studied in order to evaluate $\boldsymbol{\beta}_0$ using steps 2 and 3 of the sanity check proposed above. The probabilities of this single choice set supply a reasonable quick test of the appropriateness of $\boldsymbol{\beta}_0$.

From the above discussion on the sanity check for the prior mean $\boldsymbol{\beta}_0$, it is clear that the number of attributes plays a role in the specification of $\boldsymbol{\beta}_0$. The more attributes are involved, the more extreme the logit probabilities for any choice set might be. In particular, the probabilities for the choice set with the most extreme alternatives might be close to zero and one. The more extreme the probabilities, the more confident one is supposed to be about the preferences for the alternatives. Consequently, the probabilities may readily overstate one's beliefs. In case of a large number of attributes, we therefore advise against taking a prior mean far away from the zero vector

and recommend using smaller absolute prior parameter values.

We illustrate this argument by comparing the prior means $\boldsymbol{\beta}_{01} = [-1, -1]'$ and $\boldsymbol{\beta}_{02} = [-1, -1, -1, -1, -1, -1]'$ associated with two and six two-level attributes, respectively. Both prior means assume equally spaced elements between -1 and 1 for the levels of each attribute. The choice sets with the most extreme alternatives given each of these priors are shown in Tables 3.2a and 3.2b. Using $\boldsymbol{\beta}_{01}$ the probabilities that Alternatives I and II in Table 3.2a are chosen are 0.01799 and 0.98201, respectively. Using $\boldsymbol{\beta}_{02}$ the probabilities that Alternatives I and II in Table 3.2b are chosen equal 0.00001 and 0.99999.

Table 3.2: Two choice sets with the most extreme alternatives given a) $\boldsymbol{\beta}_{01} = [-1, -1]'$ and b) $\boldsymbol{\beta}_{02} = [-1, -1, -1, -1, -1, -1]'$.

a)

Alt	Attr	
	1	2
I	1	1
II	2	2

b)

Alt	Attributes					
	1	2	3	4	5	6
I	1	1	1	1	1	1
II	2	2	2	2	2	2

It is obvious that the probabilities for the choice set with two attributes in Table 3.2a are less extreme than those for the choice set with six attributes in Table 3.2b. For the choice set in Table 3.2b, one has to be virtually certain about the alternative that people prefer, whereas for the choice set in Table 3.2a, there is still room for a little hesitation. We believe that, without any data from a previous study, it is very rare to be completely confident about people's preference evaluations for the choice set in Table 3.2b.

In fact, already in the case of three attributes, the most extreme logit probabilities, being 0.00247 and 0.99753, most probably overvalue one's notions. Note that these probabilities are independent of the number of levels for each attribute if the parameter values in $\boldsymbol{\beta}_0$ are evenly spaced between -1 and 1 per attribute. Only the value of -1 for the first level is important for each attribute then since the values for the other levels cancel each other out (for an example see Section 3.3.2). So we do not advocate the use of a prior mean

$\boldsymbol{\beta}_0$ with equally spaced elements between -1 and 1 for each attribute in the case of more than two attributes either.

Concerning the specification of the prior variance-covariance matrix $\boldsymbol{\Sigma}_0$, we argue that the variances should not be larger than 1. This is because a prior variance of 1 already indicates a great amount of uncertainty.

3.3 Bayesian designs outperforming utility-neutral designs

We now show with two design cases how Bayesian optimal designs outperform utility-neutral optimal designs on average. We focus on Bayesian designs computed by means of the $\mathcal{D}_B$- and $\mathcal{V}_B$-optimality criteria since these are the most appealing criteria from an estimation and prediction viewpoint, respectively (Kessels et al. 2006). Moreover, Kessels, Goos and Vandebroek (2006) demonstrated that the $\mathcal{D}_B$-optimality criterion also scores well in terms of prediction. In both design cases, we assume main-effects models for which no prior information is available from previous studies.

3.3.1 The $2^6/2/8$ case: $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal choice designs versus an orthogonally blocked fractional factorial design

For a first design case, we computed $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs and a utility-neutral optimal design of class $2^6/2/8$. The design profiles are thereby described by six two-level attributes and are grouped two by two in each of eight choice sets. So in total, the designs consist of 16 profiles each. Using effects-type coding, the number of elements, k, in the parameter vector is 6.

We constructed the Bayesian designs by assuming some prior beliefs about people's preferences for the attribute levels. In accordance with the guidelines presented in Section 3.2, we incorporated these beliefs in the prior parameter distribution $\pi(\boldsymbol{\beta}) = \mathcal{N}(\boldsymbol{\beta}|\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$ by specifying the prior mean as

$\boldsymbol{\beta}_0 = [-0.5, -0.5, -0.5, -0.5, -0.5, -0.5]'$ and the prior variance-covariance matrix as $\boldsymbol{\Sigma}_0 = 0.7^2 \times \mathbf{I}_6$. We created the $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs using the adaptive algorithm of Kessels et al. (2006) provided in MATLAB 7. We performed $1,000$ tries or random starts of this algorithm for each of the criteria. As input to the algorithm, we constructed a systematic 20-point sample for generating the tries and drew a random $1,000$-point sample for evaluating the resulting designs.

As a utility-neutral optimal design, we used an orthogonally blocked 2^{6-2} fractional factorial design with blocks of size two. This fractional factorial design is locally $\mathcal{D}$-, $\mathcal{A}$-, $\mathcal{G}$- and $\mathcal{V}$-optimal for $\boldsymbol{\beta}_P = [0, 0, 0, 0, 0, 0]'$ given the present choice design configuration. We produced it in JMP 6. The orthogonally blocked 2^{6-2} fractional factorial design and the $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs appear in Table 3.3. As can be seen, the choice sets of the fractional factorial design are completely level balanced, whereas those of the Bayesian optimal designs exhibit some level overlap.

Figure 3.1 contains two plots comparing the utility-neutral optimal design or the orthogonally blocked 2^{6-2} fractional factorial design to the $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs. Figure 3.1a shows the relative $\mathcal{D}_P$-efficiencies of the fractional factorial design to the $\mathcal{D}_B$-optimal design for various true parameter vectors and Figure 3.1b shows the relative $\mathcal{V}_P$-efficiencies of the fractional factorial design to the $\mathcal{V}_B$-optimal design. The plusses in the graphs correspond to true parameter vectors going from $[-1.5, -1.5, -1.5, -1.5, -1.5, -1.5]'$ through the prior mean of the Bayesian designs, $\boldsymbol{\beta}_0 = [-0.5, -0.5, -0.5, -0.5, -0.5, -0.5,]'$, and finally to the implied prior mean of the utility-neutral design, $\boldsymbol{\beta}_P = [0, 0, 0, 0, 0, 0]'$. Thus each plus sign represents a true parameter of the form $[c, c, c, c, c, c]'$ where c is on the interval $[-1.5\ 0]$.

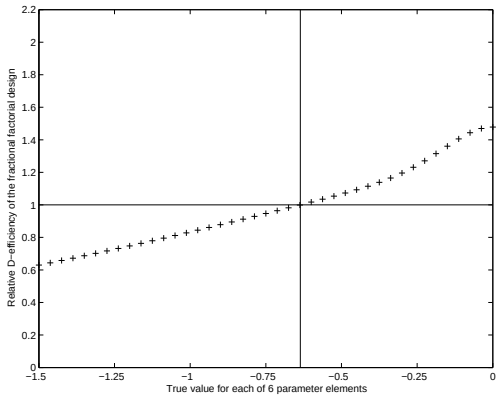
At the far left hand side of Figure 3.1a comparing $\mathcal{D}_P$-efficiencies, the $\mathcal{D}_B$-optimal design is about 40% more efficient than the utility-neutral design. The relative $\mathcal{D}_P$-efficiency of the utility-neutral design increases until $c = -0.64$ where the two designs are equally efficient. For less negative values of c, the utility-neutral design is more efficient than the $\mathcal{D}_B$-optimal design. Consequently, at the prior mean of the Bayesian designs, where $c = -0.5$, the

Table 3.3: An orthogonally blocked 2^{6-2} fractional factorial design used as utility-neutral optimal design and the $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs for the $2^6/2/8$ example.

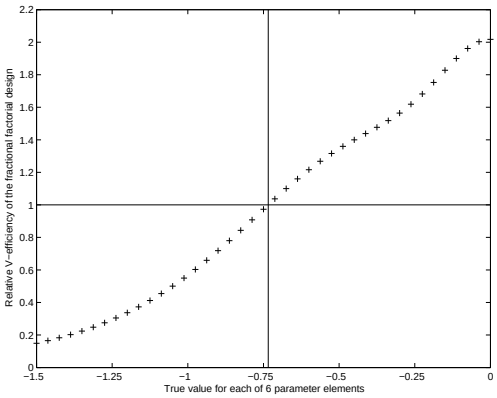
		2^{6-2} FracF						$\mathcal{D}_B$						$\mathcal{V}_B$					
Choice	Alt	Attr						Attr						Attr					
set		1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
1	I	2	2	2	1	1	1	1	1	2	1	2	1	2	1	1	1	1	2
	II	1	1	1	2	2	2	1	1	1	2	1	2	1	1	1	1	2	2
2	I	2	1	1	2	2	1	2	2	2	1	2	2	1	1	2	2	1	2
	II	1	2	2	1	1	2	1	1	2	2	2	2	2	1	1	1	2	2
3	I	1	2	1	1	2	1	2	1	1	1	1	2	2	1	1	2	1	2
	II	2	1	2	2	1	2	1	2	1	2	2	1	2	2	1	1	1	1
4	I	2	2	1	1	2	2	2	1	1	1	2	2	2	1	2	1	2	1
	II	1	1	2	2	1	1	1	2	2	1	1	1	1	2	1	1	2	2
5	I	2	2	1	2	1	1	2	1	2	2	1	1	2	2	2	1	1	2
	II	1	1	2	1	2	2	1	2	1	1	2	2	2	1	1	2	2	1
6	I	2	1	2	1	2	1	2	2	1	2	1	1	1	1	2	1	2	1
	II	1	2	1	2	1	2	1	1	2	1	1	2	2	2	1	2	2	2
7	I	2	1	1	1	1	2	2	1	1	2	2	1	2	1	1	1	1	2
	II	1	2	2	2	2	1	2	2	2	2	1	2	2	1	2	1	2	2
8	I	2	2	2	2	2	2	2	1	2	2	2	2	2	1	2	1	1	2
	II	1	1	1	1	1	1	2	2	1	1	1	2	2	2	2	2	1	1

utility-neutral design outperforms the $\mathcal{D}_B$-optimal design, but only slightly by less than 10%. For the zero parameter vector the utility-neutral design is about 45% more efficient than the $\mathcal{D}_B$-optimal design.

Figure 3.1b shows a similar trend for the relative $\mathcal{V}_P$-efficiencies. There, the crossover point for the $\mathcal{V}_B$-optimal design and the utility-neutral design to be equally efficient is found at $c = -0.73$. At the prior mean of the Bayesian designs, at $c = -0.5$, the utility-neutral design is about 35% more efficient than the $\mathcal{V}_B$-optimal design. Note, however, that at the far left hand side of this plot, the relative $\mathcal{V}_P$-efficiency of the utility-neutral design is less than 20%. Alternatively, one could say that the $\mathcal{V}_B$-optimal design is roughly five



(a) Relative $\mathcal{D}_P$-efficiencies of the 2^{6-2} design to the $\mathcal{D}_B$-optimal design



(b) Relative $\mathcal{V}_P$-efficiencies of the 2^{6-2} design to the $\mathcal{V}_B$-optimal design

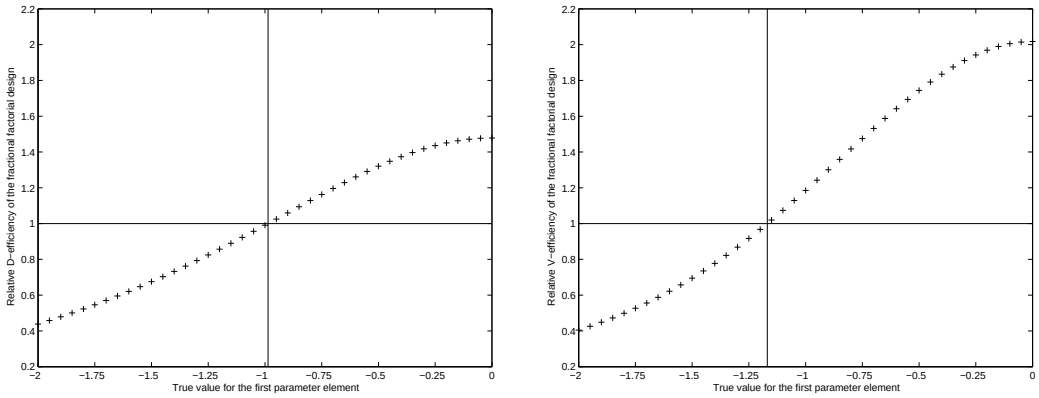
Figure 3.1: Relative local efficiencies of the orthogonally blocked 2^{6-2} fractional factorial design to the Bayesian optimal designs for various true parameter vectors starting from $[-1.5, -1.5, -1.5, -1.5, -1.5, -1.5]'$ and moving toward $[0, 0, 0, 0, 0, 0]'$ with equal values for each parameter element.

times more efficient than the utility-neutral design for the parameter vectors in this corner. By contrast, at the zero parameter vector the utility-neutral design is only twice as efficient.

In summary, we can conclude the following from Figure 3.1. While the utility-neutral optimal design is more efficient with respect to the $\mathcal{D}_P$- and $\mathcal{V}_P$-optimality criteria than the Bayesian optimal designs for true parameter vectors that are small in magnitude, the Bayesian designs are far more robust to true parameter values that are some distance away from the prior mean. Since the prior mean of the Bayesian designs has its parameter values of -0.5 fairly close to zero, the utility-neutral design is slightly more efficient there than the Bayesian designs.

To further illustrate this, we plotted similar graphs as in Figure 3.1 but for a different range of true parameter vectors. The relative $\mathcal{D}_P$-efficiencies of the fractional factorial design to the $\mathcal{D}_B$-optimal design appear in Figure 3.2a

and the relative $\mathcal{V}_P$-efficiencies of the fractional factorial design to the $\mathcal{V}_B$-optimal design appear in Figure 3.2b. Here, the true parameter vectors go from $[-2, -0.5, 0, 0, 0, 0]'$ to the implied prior mean of the utility-neutral design, $\boldsymbol{\beta}_P = [0, 0, 0, 0, 0, 0]'$. Each plus sign now corresponds to a true parameter of the form $[c, c/4, 0, 0, 0, 0]'$ where c is on the interval $[-2\ 0]$.



(a) Relative $\mathcal{D}_P$-efficiencies of the 2^{6-2} design to the $\mathcal{D}_B$-optimal design

(b) Relative $\mathcal{V}_P$-efficiencies of the 2^{6-2} design to the $\mathcal{V}_B$-optimal design

Figure 3.2: Relative local efficiencies of the orthogonally blocked 2^{6-2} fractional factorial design to the Bayesian optimal designs for various true parameter vectors starting from $[-2, -0.5, 0, 0, 0, 0]'$ and proportionally moving toward $[0, 0, 0, 0, 0, 0]'$.

At the far left hand side of Figure 3.2a, the $\mathcal{D}_B$-optimal design is more than twice as efficient as the utility-neutral design. The same can be observed for the $\mathcal{V}_B$-optimal design in terms of $\mathcal{V}_P$-efficiency at the far left hand side of Figure 3.2b. This happens despite the fact that the parameter vector $[-2, -0.5, 0, 0, 0, 0]'$ is almost equally far from $\boldsymbol{\beta}_P = [0, 0, 0, 0, 0, 0]'$ as from $\boldsymbol{\beta}_0 = [-0.5, -0.5, -0.5, -0.5, -0.5, -0.5,]'$. More specifically, if we denote $[-2, -0.5, 0, 0, 0, 0]'$ by $\boldsymbol{\beta}_t$, then the Euclidean distances $d(\boldsymbol{\beta}_t, \boldsymbol{\beta}_P)$ and $d(\boldsymbol{\beta}_t, \boldsymbol{\beta}_0)$ approximately equal two.

In Figure 3.2a the $\mathcal{D}_B$-optimal design and the utility-neutral design are equally efficient at $c = -0.99$. The crossover point for the $\mathcal{V}_B$-optimal design and

the utility-neutral design to be equally efficient occurs at $c = -1.17$ in Figure 3.2b. These values of c are more negative than in Figures 3.1a and 3.1b because four of the six parameter elements are remaining at zero which is advantageous to the utility-neutral design. Furthermore, it should be noted that many of the parameter vectors to the left of the vertical lines in Figures 3.2a and 3.2b are actually closer to the zero vector than to the prior mean of the Bayesian designs. Yet in spite of this, the $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs are more efficient than the utility-neutral design in the left panels of these figures.

We now examine more closely the estimation and prediction capabilities of the utility-neutral and Bayesian optimal designs at the true parameter $\boldsymbol{\beta}_t = [-2, -0.5, 0, 0, 0, 0]'$. In this way, we show how poorly the utility-neutral design performs when the true parameter vector consists of values at a distance away from zero. The relative $\mathcal{D}_P$- and $\mathcal{V}_P$-efficiencies of the utility-neutral and Bayesian optimal designs at the true parameter $\boldsymbol{\beta}_t$ are included in Table 3.4. The $\mathcal{D}_B$-optimal design turns out to be fairly efficient in terms of the $\mathcal{V}_P$-optimality criterion compared with the $\mathcal{V}_B$-optimal design. Also, the $\mathcal{V}_B$-optimal design is fairly efficient in terms of the $\mathcal{D}_P$-optimality criterion relatively to the $\mathcal{D}_B$-optimal design. So in terms of relative $\mathcal{D}_P$- and $\mathcal{V}_P$-efficiency at $\boldsymbol{\beta}_t$, the Bayesian designs behave similarly and contrast with the utility-neutral design.

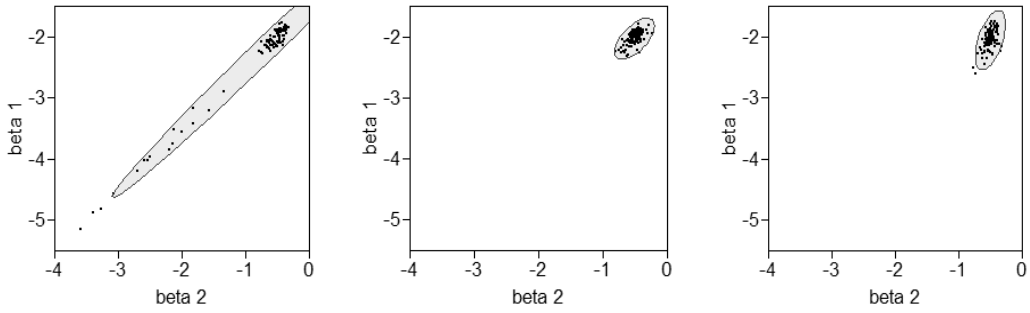
Table 3.4: Relative $\mathcal{D}_P$- and $\mathcal{V}_P$-efficiencies of the orthogonally blocked 2^{6-2} fractional factorial design and the Bayesian optimal designs to the $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs, respectively. The efficiencies are obtained at the true parameter vector $\boldsymbol{\beta}_t = [-2, -0.5, 0, 0, 0, 0]'$.

Rel. eff.	2^{6-2}	$\mathcal{D}_B$	$\mathcal{V}_B$
$\mathcal{D}_P$	44%	100%	92%
$\mathcal{V}_P$	40%	95%	100%

We further compare the estimation and prediction performance of the utility-neutral and Bayesian designs at the true parameter $\boldsymbol{\beta}_t$ with a simulation

study. Based on $\boldsymbol{\beta}_t$ we simulated 100 datasets with choices of 200 respondents for each of the fractional factorial and $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs. We subsequently estimated the parameter values for each dataset.

In Figures 3.3a, 3.3b and 3.3c we plotted the 100 estimates for $\beta_{1t} = -2$ against the 100 estimates for $\beta_{2t} = -0.5$ for the fractional factorial design and the $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs, respectively. In this way, we obtain additional information on the correlation between the estimates for β_{1t} and β_{2t}. From Figure 3.3a, we clearly observe that a substantial number of the estimates from the fractional factorial design are far away from their true values. Moreover, the estimates for β_{1t} and β_{2t} are strongly correlated. This means that if β_{1t} is poorly estimated, β_{2t} is poorly estimated as well. Not surprisingly, the estimates from the $\mathcal{D}_B$-optimal design in Figure 3.3b are all very precise, but some from the $\mathcal{V}_B$-optimal design in Figure 3.3c are less precise. For these two designs, the estimates are almost uncorrelated.



(a) Orthogonally blocked 2^{6-2} fractional factorial design

(b) $\mathcal{D}_B$-optimal design

(c) $\mathcal{V}_B$-optimal design

Figure 3.3: Scatter plots showing the correlation between 100 estimates for $\beta_{1t} = -2$ and $\beta_{2t} = -0.5$.

Figure 3.4 shows the box plots with 100 predicted probabilities based on the 100 estimates for β_{3t}, β_{4t}, β_{5t} and β_{6t} for each of the three designs. Since these four coefficients have a true value of zero, the predicted probabilities should ideally be 0.5. We thus only consider the last four attributes to study

the variability around the predicted probability of 0.5. Profiles described by these four attributes are referred to as partial profiles as they only include a subset of the attributes. Because of the zero parameter values the predicted probabilities can be calculated for any partial profile in any choice set with two partial profiles composed of the last four attributes. In the choice set we used, one partial alternative has all four attributes at the first level and the other alternative has all four attributes at the second level. We computed the predicted probabilities for the latter alternative.

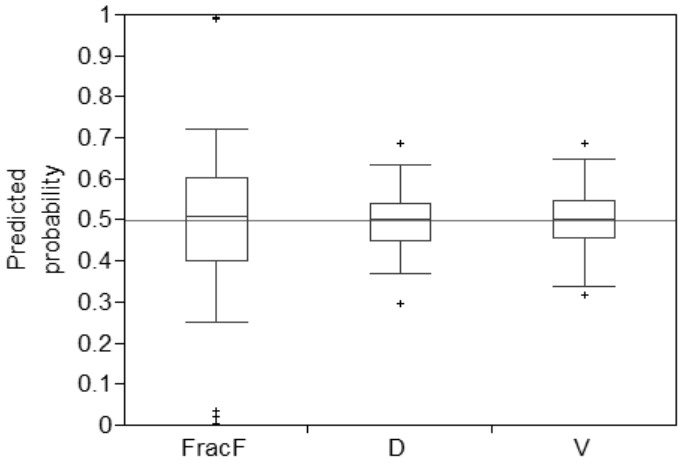


Figure 3.4: Box plots of 100 predicted probabilities based on 100 estimates for $\beta_{3t} = \beta_{4t} = \beta_{5t} = \beta_{6t} = 0$. They are shown for the orthogonally blocked 2^{6-2} fractional factorial design and the $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs.

Clearly, the box plot for the orthogonally blocked 2^{6-2} fractional factorial design is substantially wider than the box plots for the $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs. Also, there are outlying predicted probabilities near 0 and 1 for the fractional factorial design. Using Levene's test for equality of variances, the significance probability is 5×10^{-13}. As a result, there is no doubt that the predictions from the fractional factorial design have a substantially higher variance than the predictions from the $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs. Further, there is no significant difference between the quality of the predictions from the $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs.

So for a true parameter vector with one or more values reasonably large in magnitude, the relative $\mathcal{D}_P$- and $\mathcal{V}_P$-efficiencies, the scatter plots uncovering the correlation between the estimates and the box plots showing the prediction variances have all illustrated that the utility-neutral design has noticeably worse properties than the Bayesian designs. On the other hand, at the zero parameter vector, the utility-neutral design is the best design option. However, the utility-neutral design's implied prior mean of all zero values indicates that none of the attributes has much impact on consumer preferences. If this assumption were true, then it would make no sense to run the experiment. Hence, the Bayesian designs should generally be favored. They provide the best estimates and predictions on average for a whole range of true parameter vectors including true parameter values that are fairly large in magnitude.

Our conclusions so far are all based on this $2^6/2/8$ design example. In a second design example, we compare the Bayesian designs with a classical design with a completely different choice design structure. We do this to show that the characteristics we noted in the current example are not unique.

3.3.2 The $4^2/4/4$ case: $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal choice designs versus an orthogonally blocked full factorial design

In this second design case, we produced $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs and a utility-neutral optimal design of class $4^2/4/4$. Here, the profiles are configured from two attributes with 4 levels each and are arranged in 4 choice sets of size 4. As in the previous design case, the designs comprise 16 profiles. Also similar to the first case is that the number of elements, k, in the parameter vector is 6.

We generated the Bayesian designs under the assumption that one's beliefs about people's predilections are well represented by the prior parameter distribution $\pi(\boldsymbol{\beta}) = \mathcal{N}(\boldsymbol{\beta}|\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$ with $\boldsymbol{\beta}_0 = [-1, -1/3, 1/3, -1, -1/3, 1/3]'$ and $\boldsymbol{\Sigma}_0 = 0.4^2 \times \mathbf{I}_6$. As explained in Section 3.2, it is reasonable to equally space

the parameter values in $\boldsymbol{\beta}_0$ between -1 and 1 for each attribute if only two attributes are assumed. An accompanying prior variance of 0.16 thereby expresses a small amount of uncertainty. We again produced each of the $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs using $1,000$ tries of the adaptive algorithm of Kessels et al. (2006). We used a systematic sample of 20 parameters for the design generation and a Monte Carlo sample of $1,000$ parameters for the design evaluation.

For the realization of the utility-neutral optimal design, we generated an orthogonally blocked full 4^2 factorial design in JMP 6. Given the current choice design structure, this design is locally $\mathcal{D}$-, $\mathcal{A}$-, $\mathcal{G}$- and $\mathcal{V}$-optimal for $\boldsymbol{\beta}_P = [0, 0, 0, 0, 0, 0]'$. Table 3.5 shows the orthogonally blocked full 4^2 factorial design and the $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs. Like in the preceding design case, there is no level overlap in the full factorial design, but some is present in the Bayesian designs.

To demonstrate that the Bayesian optimal designs should generally be preferred to the utility-neutral optimal design or the orthogonally blocked full 4^2 factorial design, we plotted again two graphs with relative efficiencies for various true parameter vectors. They appear in Figure 3.5. Figure 3.5a shows the $\mathcal{D}_P$-efficiencies of the full factorial design relative to the $\mathcal{D}_B$-optimal design and Figure 3.5b shows the $\mathcal{V}_P$-efficiencies of the full factorial design relative to the $\mathcal{V}_B$-optimal design. The true parameter vectors go from $[-1.5, -0.5, 0.5, -1.5, -0.5, 0.5]'$ through the prior mean of the Bayesian designs, $\boldsymbol{\beta}_0 = [-1, -1/3, 1/3, -1, -1/3, 1/3]'$, to end up again at the implied prior mean of the utility-neutral design, $\boldsymbol{\beta}_P = [0, 0, 0, 0, 0, 0]'$. So each plus sign corresponds to a true parameter of the form $[c, c/3, -c/3, c, c/3, -c/3]'$ where c is on the interval $[-1.5\ 0]$.

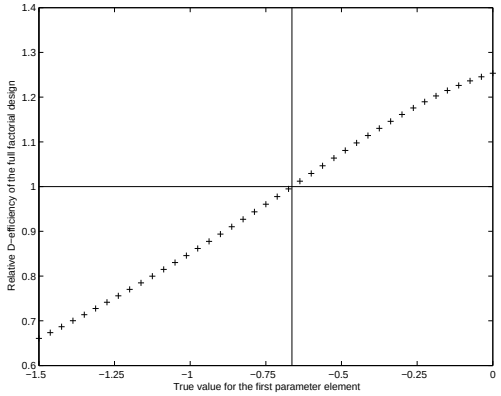
Figures 3.5a and 3.5b clearly confirm our finding that the Bayesian designs substantially outperform the utility-neutral design for parameter values reasonably large in magnitude, whereas the utility-neutral design is more efficient for parameter vectors close to the zero vector. As far as $\mathcal{D}_P$-efficiency is concerned, the far left hand side of Figure 3.5a shows that the $\mathcal{D}_B$-optimal design outperforms the utility-neutral design by approximately 35%. The

Table 3.5: An orthogonally blocked full 4^2 factorial design used as utility-neutral optimal design and the $\mathcal{D}_B$- and $\mathcal{V}_B$-optimal designs for the $4^2/4/4$ example.

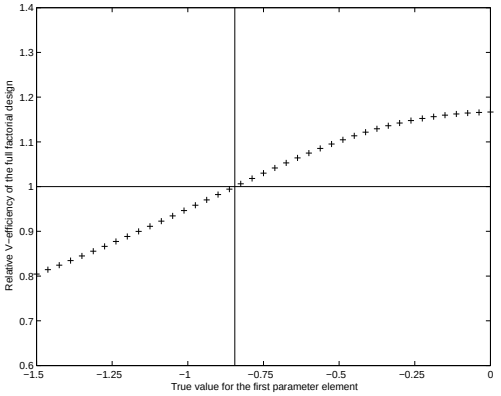
		FullF		$\mathcal{D}_B$		$\mathcal{V}_B$	
Choice	Alt	Attr		Attr		Attr	
set		1	2	1	2	1	2
1	I	2	2	4	1	3	3
	II	4	4	3	2	2	1
	III	3	3	1	3	4	2
	IV	1	1	2	1	4	1
2	I	2	3	1	4	4	3
	II	3	4	3	1	1	4
	III	1	2	1	2	3	1
	IV	4	1	2	3	2	2
3	I	1	4	3	3	1	2
	II	2	1	2	1	1	4
	III	3	2	1	2	3	4
	IV	4	3	2	2	2	3
4	I	2	4	3	1	4	1
	II	1	3	2	4	1	3
	III	4	2	4	2	3	2
	IV	3	1	1	3	2	4

efficiency gap steadily decreases until $c = -0.66$, where the two designs are equally efficient, after which it increases in favor of the utility-neutral design. At the zero parameter vector the utility-neutral design is 25% more efficient than the $\mathcal{D}_B$-optimal design. A similar course can be observed for the relative $\mathcal{V}_P$-efficiency in Figure 3.5b. Here, however, the efficiency gaps at the outer sides of the plot are smaller and the $\mathcal{V}_B$-optimal design and the utility-neutral design are equally efficient at $c = -0.84$.

The crossover points for each of the Bayesian designs and the utility-neutral design to be equally efficient are clearly larger than $c = -1$. This is because



(a) Relative $\mathcal{D}_P$-efficiencies of the 4^2 design to the $\mathcal{D}_B$-optimal design



(b) Relative $\mathcal{V}_P$-efficiencies of the 4^2 design to the $\mathcal{V}_B$-optimal design

Figure 3.5: Relative local efficiencies of the orthogonally blocked full 4^2 factorial design to the Bayesian optimal designs for various true parameter vectors starting from $[-1.5, -0.5, 0.5, -1.5, -0.5,\ 0.5]'$ and proportionally moving toward $[0, 0, 0, 0, 0, 0]'$.

the starting true parameter vector at $c = -1.5$ does not include any zero values and thereby lies relatively far from the zero vector. We could also observe this in Figures 3.1a and 3.1b. Consequently, at the prior mean of the Bayesian designs, where $c = -1$, the Bayesian designs are more efficient than the utility-neutral design. We expected this result because the prior mean lies rather far from the zero vector. Recall that in Figures 3.1a and 3.1b on the other hand, the utility-neutral design is more efficient at the prior mean of the Bayesian designs since the parameter values of -0.5 are fairly close to zero.

3.4 Conclusion

In this chapter, we had two goals. First, we wanted to provide some practical recommendations on how to properly specify the prior parameter distribution for constructing Bayesian choice designs. We did this because some of the prior distributions used in the literature are internally inconsistent. Second,

we wished to illustrate that Bayesian designs have generally better properties than utility-neutral designs. We therefore used two separate examples.

In the Bayesian choice design literature, we noticed that in the absence of prior information from a previous enquiry, the specifications of the prior mean and variance conflict. One has to be careful not to take a prior mean that is too informative compared with a specific prior variance. Therefore, we established a sanity check for the prior mean. It is built around the principle that one's expectations about the preferences for alternatives in choice sets of size two should be in line with the logit probabilities for those alternatives given the prior mean. A quick look at the choice set with the most extreme alternatives already provides some profound insights about the prior mean's suitability. Furthermore, we advise to take a prior variance of one as upper limit for the specification of the variances as this indicates already a lot of uncertainty.

In the choice design literature, the Bayesian design approach competes with the linear design approach for the production of choice designs. The Bayesian approach should however be favored because Bayesian designs are constructed for a prior parameter distribution incorporating all prior knowledge, whereas linear or utility-neutral designs are generated under the assumption that all alternatives are equally preferred by the respondents. Utility-neutral designs can thus be seen as Bayesian designs with zero prior mean and prior variance. Note that even if the prior variance is very small around zero, Bayesian designs are utility-neutral designs. However, if one believes in these specifications behind utility-neutral designs, then it would make no sense to run the experiment. A zero prior mean is only justified if it is accompanied by a large prior variance to identify the situation where one is completely without intuition about people's preferences. In that case, Bayesian designs differ from utility-neutral designs.

Not surprisingly therefore, our study of two design examples showed that Bayesian designs substantially outperform utility-neutral designs whenever some true parameter values are reasonably large in magnitude, whereas utility-neutral designs are more efficient for true parameter vectors close to

the zero vector. As one generally conducts an experiment when one anticipates a number of important attribute levels, and thus a number of fairly large parameter values, Bayesian designs should clearly be preferred.

Part II

Rating-based conjoint design

Chapter 4

Rating-based conjoint designs

This chapter has been submitted as
☞ Kessels, R., Goos, P. and Vandebroek M. (2006). Optimal designs for conjoint experiments.

A shorter version has been published as
☞ Kessels, R., Goos, P. and Vandebroek M. (2005). Estimating people's values, *Scientific Computing World – ENBIS Magazine* **84**: 40–41.

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. These profiles are expensive prototypes that respondents must test in advance. 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.

4.1 Introduction

In marketing, conjoint experiments have frequently been carried out to measure consumer preferences for the attributes of various products or services, jointly referred to as goods (Green, Krieger and Wind 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, Vriens and Burhenne 1994; Gustafsson, Herrmann and Huber 2003). In a traditional conjoint experiment, respondents are usually asked to rate a set of goods on a scale. These goods are presented as profiles or alternatives of combinations of different attribute levels. Besides rating on a scale, rating may also occur by directly asking reservation prices for the profiles. A reservation price for a good is the highest price a consumer is willing to pay for the good.

In conjoint experiments with prototypes as profiles, budgetary constraints usually force the researcher to administer only a small number of profiles to a restricted number of respondents. Also, the researcher is limited in the number of prototypes to assign to each respondent because each respondent must test these prototypes in advance. Given this experimental situation the researcher wants to elicit as much information as possible on the utilities the respondents derive from the attribute levels of the good. The utilities are also called part-worths and correspond to the parameters of a statistical model. Following an accurate estimation of these parameters, the researcher uses the model to learn about consumers' trade-offs as well as to make precise predictions about their future purchasing behavior.

Now, the quality of these inferences highly depends on the profiles and the number of test persons used in the conjoint study. This is even more so if only a small number of profiles can be used. Also the assignment of the profiles to the subjects plays a key role. As a consequence, an efficient experimental design is required. The experimental design literature on conjoint experiments is however silent about how to carefully select sets of alternative prototypes to be evaluated by the respondents. For example, if 30 prototypes can be developed from a set of many possible ones, then the literature does neither provide a tailor-made answer about how to select those 30 prototypes, nor

about the ideal number of test persons or the assignment of the 30 selected alternatives to these respondents. To provide answers to these questions is the goal of this paper.

The method we adopt to solve the conjoint design problem is based on the optimal design approach for blocked and split-plot experiments advocated by Goos and Vandebroek (2001a; 2001b; 2004) and Goos (2002; 2006a). Block designs are heavily used in industry and agriculture when not all the observations can be carried out under homogeneous circumstances, for example when more than one batch of material is required or when the experiment takes up more than one day. Split-plot designs are special cases of block designs where some of the experimental factors stay constant within each block. In all of the work on block and split-plot designs, the assumed model is the linear random block effects model. Like Brazier, Roberts and Deverill (2002) we adopt this model in the conjoint setting and refer to it as the linear random respondent effects model.

The motivation for this model is as follows. It is reasonable to assume that respondents are randomly selected from a population and that they are heterogeneous. Respondent heterogeneity is due to variations in terms of age, experience with the good under study, physical characteristics, cognitive abilities, and so forth. The consequence of this heterogeneity is that profile ratings from different respondents are more dissimilar than profile ratings from the same respondent. The likeness of the ratings from a single respondent is nothing but a positive correlation. To capture the heterogeneity between respondents, or the correlation within respondents, a random effect is included in the model.

Drawing on the random respondent effects model, a conjoint design consists of blocks or sets of profiles that are each offered to a different respondent and the number of respondents is equal to the number of blocks in the design. Note that we focus on main-effects conjoint designs only. To evaluate different conjoint designs, we apply the $\mathcal{D}$-optimality criterion that aims at designs that minimize the determinant of the variance-covariance matrix of the parameter estimators. To search for efficient conjoint designs, we use

an adapted version of the split-plot design construction algorithm of Goos and Vandebroek (2004). The adapted algorithm finds the $\mathcal{D}$-optimal number of respondents and the $\mathcal{D}$-optimal number of profiles for each respondent in addition to the $\mathcal{D}$-optimal design profiles. It is possible that the algorithm assigns sets of different sizes to the respondents since Atkinson and Donev (1992) and Goos (2002; 2006a) all provided examples of $\mathcal{D}$-optimal block or split-plot designs with heterogeneous block sizes.

Other approaches to deal with respondent heterogeneity for design construction can be found in Liski et al. (2002) and Entholzner et al. (2005). 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, Cochran and Cox (1957) recommended balanced incomplete block designs for preference rating as most of the design plans contain blocks with six or fewer units. Regrettably, these designs are only built for one qualitative factor, the levels of which are called treatments. Typical for balanced incomplete block designs is that they are universally optimal for the estimation of the treatment and the block effect. 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.

This chapter is organized as follows. First, we embark on the random respondent effects model in Section 4.2. Section 4.3 explains how to analyze a conjoint experiment assuming this model and Section 4.4 discusses the design criterion. Next, Section 4.5 presents the design construction algorithm and Section 4.6 describes the computational results. Finally, Section 4.7 summarizes the chapter and proposes future research directions.

4.2 The model

The model used to design and analyze rating-based conjoint experiments is a random respondent effects model. Suppose a conjoint experiment is set up for n profiles. In total, b respondents are appointed who each rate a different set or block of profiles so as to be able to estimate all parameters. The sizes of these profile sets may be unequal. The n profiles are accordingly arranged in b sets of sizes $m_1, ..., m_b$, where $n = \sum_{i=1}^{b} m_i$. If we assume that the respondents are heterogeneous and randomly selected from a prespecified population, then the rating U_{ij} for profile j provided by respondent i is modelled as

$$U_{ij} = \mathbf{x}_{ij}'\boldsymbol{\beta} + \gamma_i + \varepsilon_{ij}. \tag{4.1}$$

In this model, $\mathbf{x}_{ij}$ is a $(k+1) \times 1$ vector having a one as first element and the attribute levels describing profile j that is rated by respondent i as the remaining k elements. The attributes are quantitative or categorical factors that can have any number of levels. $\boldsymbol{\beta} = [\beta_0, ..., \beta_k]'$ is the $(k+1)\times 1$ unknown fixed parameter vector with β_0 the intercept and $\beta_1, ..., \beta_k$ the part-worths or weights associated with the attribute levels. They reflect the importance of the levels as viewed by the average respondent and involve main effects only. γ_i represents the random effect of respondent i and ε_{ij} is a random error term.

Note that we coded the levels of the attributes by means of effects-type coding. For example, for a two-level attribute one level is coded as 1 and the other level as -1. For a three-level attribute the codings are [1 0], [0 1] and $[-1 \ -1]$. For a four-level attribute they are [1 0 0], [0 1 0], [0 0 1] and $[-1 \ -1 \ -1]$, and so forth for higher-level attributes. Also other types of coding may be used as the choice of coding does not affect the relative design efficiency in case the $\mathcal{D}$-optimality criterion is used (see Section 4.4).

In matrix notation, model (4.1) becomes

$$\boldsymbol{U} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \tag{4.2}$$

where $\boldsymbol{U}$ is a vector of n profile ratings, 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 $n \times (k+1)$ and $n \times b$, respectively. $\mathbf{X}$ is given by

$$\mathbf{X} = [\mathbf{X}_1', ..., \mathbf{X}_b']', \tag{4.3}$$

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

$$\mathbf{Z} = \text{diag}[\mathbf{1}_{m_1}, ..., \mathbf{1}_{m_b}], \tag{4.4}$$

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

$$\text{E}(\boldsymbol{\varepsilon}) = \mathbf{0}_n \text{ and } \text{Cov}(\boldsymbol{\varepsilon}) = \sigma^2_{\boldsymbol{\varepsilon}}\mathbf{I}_n, \tag{4.5}$$

$$\text{E}(\boldsymbol{\gamma}) = \mathbf{0}_b \text{ and } \text{Cov}(\boldsymbol{\gamma}) = \sigma^2_{\boldsymbol{\gamma}}\mathbf{I}_b, \tag{4.6}$$

$$\text{and } \text{Cov}(\boldsymbol{\gamma}, \boldsymbol{\varepsilon}) = \mathbf{0}_{b \times n}, \tag{4.7}$$

where $\sigma^2_{\boldsymbol{\varepsilon}}$ is the within-respondents variance and $\sigma^2_{\boldsymbol{\gamma}}$ is the between-respondents variance. Under these assumptions, the variance-covariance matrix $\mathbf{V}$ of the profile ratings $\boldsymbol{U}$ can be written as

$$\begin{aligned}
\mathbf{V} &= \text{Cov}(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}), \\
&= \text{Cov}(\mathbf{Z}\boldsymbol{\gamma}) + \text{Cov}(\boldsymbol{\varepsilon}), \\
&= \sigma^2_{\boldsymbol{\gamma}}\mathbf{Z}\mathbf{Z}' + \sigma^2_{\boldsymbol{\varepsilon}}\mathbf{I}_n, \\
&= \sigma^2_{\boldsymbol{\gamma}}\text{diag}[\mathbf{1}_{m_1}\mathbf{1}_{m_1}', ..., \mathbf{1}_{m_b}\mathbf{1}_{m_b}'] + \sigma^2_{\boldsymbol{\varepsilon}}\text{diag}[\mathbf{I}_{m_1}, ..., \mathbf{I}_{m_b}], \\
&= \text{diag}[\sigma^2_{\boldsymbol{\varepsilon}}\mathbf{I}_{m_1} + \sigma^2_{\boldsymbol{\gamma}}\mathbf{1}_{m_1}\mathbf{1}_{m_1}', ..., \sigma^2_{\boldsymbol{\varepsilon}}\mathbf{I}_{m_b} + \sigma^2_{\boldsymbol{\gamma}}\mathbf{1}_{m_b}\mathbf{1}_{m_b}'].
\end{aligned} \tag{4.8}$$

Substituting

$$\mathbf{V}_i = \sigma^2_{\boldsymbol{\varepsilon}}\mathbf{I}_{m_i} + \sigma^2_{\boldsymbol{\gamma}}\mathbf{1}_{m_i}\mathbf{1}_{m_i}', \tag{4.9}$$

leads to the variance-covariance matrix

$$\mathbf{V} = \text{diag}[\mathbf{V}_1, ..., \mathbf{V}_b]. \tag{4.10}$$

Note that the 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 from two respondents, one of whom

rated a block of two profiles and the other a block of three profiles, is

$$\mathbf{V} = \begin{bmatrix} \sigma_{\varepsilon}^2 + \sigma_{\gamma}^2 & \sigma_{\gamma}^2 & 0 & 0 & 0 \\ \sigma_{\gamma}^2 & \sigma_{\varepsilon}^2 + \sigma_{\gamma}^2 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{\varepsilon}^2 + \sigma_{\gamma}^2 & \sigma_{\gamma}^2 & \sigma_{\gamma}^2 \\ 0 & 0 & \sigma_{\gamma}^2 & \sigma_{\varepsilon}^2 + \sigma_{\gamma}^2 & \sigma_{\gamma}^2 \\ 0 & 0 & \sigma_{\gamma}^2 & \sigma_{\gamma}^2 & \sigma_{\varepsilon}^2 + \sigma_{\gamma}^2 \end{bmatrix}. \tag{4.11}$$

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

$$\rho = \frac{\sigma_{\gamma}^2}{\sigma_{\varepsilon}^2 + \sigma_{\gamma}^2}. \tag{4.12}$$

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 $\sigma_{\gamma}^2 \to 0$, the profile ratings from the same respondent are no longer correlated. In that case, $\gamma_1 = ... = \gamma_b = 0$ and as a result, the random respondent effects model (4.2) degenerates to the uncorrelated model

$$\boldsymbol{U} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}. \tag{4.13}$$

4.3 Analysis

If the error terms and the respondent effects are normally distributed, the maximum likelihood estimator of the unknown fixed parameter vector $\boldsymbol{\beta}$ in (4.1) and (4.2) is the generalized least squares (GLS) estimator. The GLS estimator is the best linear unbiased estimator (BLUE) and is given by

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\,\boldsymbol{U}, \tag{4.14}$$

with variance-covariance matrix

$$\mathrm{Cov}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}. \tag{4.15}$$

Sometimes, the variance components $\sigma^2_{\boldsymbol{\gamma}}$ and $\sigma^2_{\boldsymbol{\varepsilon}}$ 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 $\sigma^2_{\boldsymbol{\gamma}}$ and $\sigma^2_{\boldsymbol{\varepsilon}}$ are unknown and therefore, (4.14) and (4.15) cannot be applied directly. Instead, the variance components $\sigma^2_{\boldsymbol{\gamma}}$ and $\sigma^2_{\boldsymbol{\varepsilon}}$ have to be estimated, for example via restricted maximum likelihood (REML) (Gilmour and Trinca 2000). The estimates $\hat{\sigma}^2_{\boldsymbol{\gamma}}$ and $\hat{\sigma}^2_{\boldsymbol{\varepsilon}}$ are then substituted in the GLS estimator (4.14), yielding the feasible GLS estimator

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\mathbf{V}}^{-1}\,\boldsymbol{U}, \tag{4.16}$$

where

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

In that case, the variance-covariance matrix (4.15) can be approximated by

$$\mathrm{Cov}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})^{-1}. \tag{4.17}$$

4.4 Design criterion

In this chapter, we evaluate alternative conjoint design options by means of the $\mathcal{D}$-optimality criterion. The $\mathcal{D}$-optimality criterion seeks designs that minimize the determinant of the variance-covariance matrix (4.15), or equivalently, that maximize the determinant of the information matrix $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$. $\mathcal{D}$-optimal conjoint designs therefore minimize the generalized variance of the parameter estimators (Atkinson and Donev 1992). Goos and Vandebroek (2001b) showed that, because of the compound symmetric error structure of model (4.2), the information matrix is equal to

$$\mathbf{X}'\mathbf{V}^{-1}\mathbf{X} = \frac{1}{\sigma^2_{\boldsymbol{\varepsilon}}}\left\{\mathbf{X}'\mathbf{X} - \sum_{i=1}^{b}\frac{\sigma^2_{\boldsymbol{\gamma}}/\sigma^2_{\boldsymbol{\varepsilon}}}{1+m_i(\sigma^2_{\boldsymbol{\gamma}}/\sigma^2_{\boldsymbol{\varepsilon}})}(\mathbf{X}_i'\mathbf{1}_{m_i})(\mathbf{X}_i'\mathbf{1}_{m_i})'\right\}. \tag{4.18}$$

In terms of the coefficient of correlation (4.12), (4.18) can be rewritten as

$$\mathbf{X}'\mathbf{V}^{-1}\mathbf{X} = \frac{1}{\sigma^2_{\boldsymbol{\varepsilon}}}\left\{\mathbf{X}'\mathbf{X} - \sum_{i=1}^{b}\frac{\rho}{1+\rho(m_i-1)}(\mathbf{X}_i'\mathbf{1}_{m_i})(\mathbf{X}_i'\mathbf{1}_{m_i})'\right\}. \tag{4.19}$$

If $\rho \to 0$, the information matrix (4.19) reduces to

$$\frac{1}{\sigma_{\varepsilon}^2}\mathbf{X}'\mathbf{X}, \tag{4.20}$$

which is the information matrix for the uncorrelated model (4.13) that is used for analyzing data from a completely randomized experiment. Because the respondents are homogeneous in that case, the grouping of profiles in sets is irrelevant. We therefore label 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). On the other hand, we call a design that maximizes the $\mathcal{D}$-criterion value $|\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}|$ for ρ differing from zero 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 $\rho \to 1$, the information matrix (4.19) is 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, this finding implies that 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 4.6, we show that $\mathcal{D}$-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 (4.20), the $\mathcal{D}$-criterion value of a $\mathcal{D}$-optimal CRD becomes

$$|\mathbf{X}'\mathbf{X}|. \tag{4.21}$$

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

$$(1-\rho)^{-(k+1)}|\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}|, \tag{4.22}$$

with $k+1$ 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

$$|\mathbf{X}'\mathbf{X}|^{1/(k+1)} \text{ and } (1-\rho)^{-1}|\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}|^{1/(k+1)}. \tag{4.23}$$

4.5 Conjoint design algorithm

The algorithm to construct $\mathcal{D}$-optimal designs for conjoint experiments is an adaptation from that of Goos and Vandebroek (2004) for the production of $\mathcal{D}$-optimal split-plot designs allowing for variable block sizes. Given the sample size n and the degree of correlation ρ the conjoint design algorithm finds the $\mathcal{D}$-optimal number of respondents, the $\mathcal{D}$-optimal number of profiles for each respondent as well as the $\mathcal{D}$-optimal design profiles.

The algorithm starts with the composition of the set of candidate profiles. For example, for one type of designs studied in Section 4.6, we used four attributes, each acting at three levels so that the candidate set 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 candidate set. 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 with the largest prediction variance. Also in that case, the profiles are randomly assigned to an existing or to a new block.

To improve the starting design, two procedures are applied consecutively, namely interchanging design profiles from different blocks and exchanging design profiles with candidate profiles. The main procedure is the exchange procedure adjusted from the algorithm of Goos and Vandebroek (2004). The additional interchange procedure is meant to strengthen the conjoint design algorithm. The reason is that the construction of conjoint designs is more involved than the construction of split-plot designs because the levels of all attributes are allowed to vary in the blocks of conjoint designs. The interchange procedure is similar to the one in the algorithm of Goos and Vandebroek (2001a) for the construction of $\mathcal{D}$-optimal designs with given block sizes in the presence of random block effects.

In the interchange procedure of the conjoint design algorithm, 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 strategies are considered for each combination of a design profile and a candidate profile. In each of the strategies, the design profile is removed from the design and the candidate profile is added to the design. First, the candidate profile 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 conjoint design algorithm subsequently 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. Obviously, the more observations, attributes and attribute levels are involved, the more designs are possible and the higher the chance the algorithm yields local optima that are far from the global optimum. Therefore, for large problem situations, a great number of tries is needed. For example, we used $6,000$ tries to com-

pute $\mathcal{D}$-optimal conjoint designs consisting of 60 profiles in the case of four attributes, each acting at three levels.

To benchmark the results of our algorithm, we used the algorithm of Goos and Vandebroek (2001a) to which we refer as the fixed block size algorithm because it restricts its searches to conjoint designs with predetermined block sizes. Whereas the conjoint design algorithm generates the $\mathcal{D}$-optimal number of respondents, b, and the $\mathcal{D}$-optimal number of profiles for each respondent, $m_1, ..., m_b$, the fixed block size algorithm requires the values b and $m_1, ..., m_b$ as an input. The fact that the fixed block size algorithm constrains the design structure by requiring the specification of b and $m_1, ..., m_b$ means that the resulting designs do not necessarily have the optimal blocking structure and therefore may not be as efficient as possible. The conjoint design algorithm relaxes this restriction so that it does find the optimal design structure.

Like the conjoint design algorithm, the fixed block size algorithm consists of an interchange and exchange procedure. The exchange procedure is naturally limited to replacing a design profile by a candidate profile in the same block. Moreover, the design space the fixed block algorithm has to explore is much more constrained than that of the conjoint design algorithm. As a result, the fixed block size algorithm requires fewer tries and less computing time.

In the next section, we show that the $\mathcal{D}$-optimal conjoint designs are not very sensitive to ρ so that a rough estimate usually suffices. Often, a reasonable value for ρ is 0.5. For example, a dataset collected in a health economics study by Brazier, Roberts and Deverill (2002) yielded an estimate for ρ of 0.62. Also, four datasets from sensory experiments performed by the multinational brewer InBev yielded estimates for ρ of 0.48, 0.46, 0.36 and 0.41.

4.6 Results

We now present a selection of computational results from which we derive recommendations to produce $\mathcal{D}$-optimal conjoint designs when the profile construction is expensive. We first show that it is statistically justified to apply these designs instead of $\mathcal{D}$-optimal CRDs. We then proceed with a discussion of the $\mathcal{D}$-optimal blocking structures and the computing times needed. Also, we deal with some practical issues and seek ways to save computing time.

4.6.1 Designs under investigation

For our study we computed $\mathcal{D}$-optimal conjoint designs for four scenarios each involving four attributes. The first scenario has all four attributes at three levels. We denote this scenario as $(3, 3, 3, 3)$. The next three scenarios possess increasing amounts of heterogeneity in the numbers of attribute levels. The second scenario has the first attribute at two levels, the next two attributes at three levels and the fourth attribute at four levels. We refer to it as $(2, 3, 3, 4)$. The third scenario is similar to the second one except for the fourth attribute which has five levels: $(2, 3, 3, 5)$. The fourth scenario is entirely heterogeneous with all four attributes at a different level: $(2, 3, 4, 5)$.

Table 4.1 contains further information about the setup of our conjoint design study. For each scenario we indicated the number of candidate profiles and the number of elements, $k+1$, in $\boldsymbol{\beta}$ after coding the levels. We also specified the sample sizes, n, of the conjoint designs together with the degrees of correlation, ρ. Note that we included $\rho = 0$ to compute the $\mathcal{D}$-optimal CRDs. Also note that we considered more sample sizes in the $(3, 3, 3, 3)$ and $(2, 3, 3, 4)$ scenarios to perform some additional studies on these cases. These are described in Sections 4.6.5, 4.6.6 and 4.6.7.

Table 4.1: Setup of the conjoint design study.

Scenario	# cand.	$k+1$	n	ρ
$(3,3,3,3)$	81	9	20, 24, 30, 36, 40, 50, 60, 70, 72, 81	$0, 0.1, ..., 0.9$
$(2,3,3,4)$	72			
$(2,3,3,5)$	90	10	20, 30, 40, 50, 60, 70	
$(2,3,4,5)$	120	11	20, 30, 40, 50, 60, 70	

4.6.2 $\mathcal{D}$-criterion values of the $\mathcal{D}$-optimal conjoint designs and $\mathcal{D}$-optimal CRDs

The $\mathcal{D}$-criterion values of the $\mathcal{D}$-optimal conjoint designs and $\mathcal{D}$-optimal CRDs for the four scenarios appear in Table A1 of Appendix A. As we discussed in Section 4.3, we calculated these values using the expressions (4.23) to compare the $\mathcal{D}$-optimal conjoint designs with the $\mathcal{D}$-optimal CRDs for the same sample size and scenario. To gain better insight in the $\mathcal{D}$-criterion values associated with each scenario, we plotted them against the sample size for the different degrees of correlation. Figure 4.1 contains the graph for the $(3,3,3,3)$ scenario. The plots for the other scenarios show a similar trend and have therefore been left out.

It turns out that the $\mathcal{D}$-optimal CRDs are outperformed by each of the corresponding $\mathcal{D}$-optimal conjoint designs. We expected this result as Goos (2002, page 133) proved that $\mathcal{D}$-optimal block designs are more efficient than $\mathcal{D}$-optimal CRDs provided the experimental situation exhibits a block format. 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. Moreover, Figure 4.1 clearly shows 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. Figure 4.1 further reveals that the $\mathcal{D}$-criterion values increase linearly with the sample size and do not saturate after a certain number of observations.

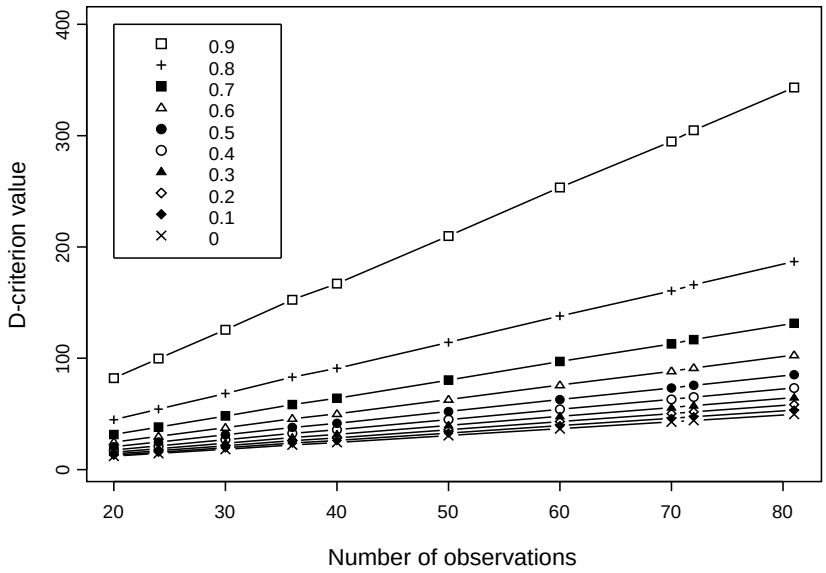


Figure 4.1: $\mathcal{D}$-criterion values of the $\mathcal{D}$-optimal conjoint designs ($\rho \neq 0$) and $\mathcal{D}$-optimal CRDs ($\rho = 0$) for the $(3, 3, 3, 3)$ scenario.

4.6.3 $\mathcal{D}$-optimal blocking structures

Because the $\mathcal{D}$-optimal conjoint designs are statistically more efficient than the $\mathcal{D}$-optimal CRDs, it is better to have a select number of respondents evaluate several profiles. In that case, the random respondent effects model (4.2) is appropriate. We now examine the $\mathcal{D}$-optimal conjoint designs to derive more precisely what the optimal number of respondents is for a specific conjoint setting, what the optimal number of profiles is for each of them to rate, and what the optimal profiles are. Table 4.2 contains the blocking structures of the $\mathcal{D}$-optimal conjoint designs for the $(3, 3, 3, 3)$ and $(2, 3, 3, 4)$ scenarios. The blocking structures pertaining to the $(2, 3, 3, 5)$ and $(2, 3, 4, 5)$ scenarios appear in Tables 4.3 and 4.4, respectively. The designs themselves are not shown, but can be obtained from the authors.

Table 4.2: Blocking structures of the $\mathcal{D}$-optimal conjoint designs for the $(3,3,3,3)$ and $(2,3,3,4)$ scenarios.

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

Table 4.3: Blocking structures of the $\mathcal{D}$-optimal conjoint designs for the $(2, 3, 3, 5)$ scenario.

Design	n	ρ	Blocking structure	b
1	20	$\{0.1; 0.2\}$	$m_1, ..., m_4 = 3 \mid m_5, m_6 = 4$	6
2	20	$\{0.3; 0.4; 0.5\}$	$m_1, m_2 = 3 \mid m_3, m_4 = 5 \mid m_5 = 4$	5
3	20	$\{0.6; ...; 0.9\}$	$m_1, ..., m_4 = 5$	4
4	30	$\{0.1\}$	$m_1, ..., m_{10} = 3$	10
5	30	$\{0.2; 0.3; 0.4\}$	$m_1, ..., m_5 = 3 \mid m_6, m_7, m_8 = 5$	8
6	30	$\{0.5; ...; 0.9\}$	$m_1, ..., m_6 = 5$	6
7	40	$\{0.1; 0.2\}$	$m_1, ..., m_{10} = 3 \mid m_{11}, m_{12} = 5$	12
8	40	$\{0.3; ...; 0.8\}$	$m_1, ..., m_5 = 3 \mid m_6, ..., m_{10} = 5$	10
9	40	$\{0.9\}$	$m_1, ..., m_8 = 5$	8
10	50	$\{0.1; 0.2; 0.3\}$	$m_1, ..., m_{10} = 3 \mid m_{11}, ..., m_{14} = 5$	14
11	50	$\{0.4; ...; 0.7\}$	$m_1, ..., m_5 = 3 \mid m_6, ..., m_{12} = 5$	12
12	50	$\{0.8; 0.9\}$	$m_1, ..., m_{10} = 5$	10
13	60	$\{0.1\}$	$m_1, ..., m_{20} = 3$	20
14	60	$\{0.2\}$	$m_1, ..., m_{15} = 3 \mid m_{16}, m_{17}, m_{18} = 5$	18
15	60	$\{0.3; 0.4\}$	$m_1, ..., m_{10} = 3 \mid m_{11}, ..., m_{16} = 5$	16
16	60	$\{0.5; ...; 0.9\}$	$m_1, ..., m_5 = 3 \mid m_6, ..., m_{14} = 5$	14
17	70	$\{0.1\}$	$m_1, ..., m_{20} = 3 \mid m_{21}, m_{22} = 5$	22
18	70	$\{0.2\}$	$m_1, ..., m_{15} = 3 \mid m_{16}, ..., m_{20} = 5$	20
19	70	$\{0.3; 0.4\}$	$m_1, ..., m_{10} = 3 \mid m_{11}, ..., m_{18} = 5$	18
20	70	$\{0.5; ...; 0.9\}$	$m_1, ..., m_5 = 3 \mid m_6, ..., m_{16} = 5$	16

Table 4.4: Blocking structures of the $\mathcal{D}$-optimal conjoint designs for the $(2, 3, 4, 5)$ scenario.

Design	n	ρ	Blocking structure	b
1	20	$\{0.1; ...; 0.9\}$	$m_1, ..., m_5 = 4$	5
2	30	$\{0.1\}$	$m_1, ..., m_6 = 4 \mid m_7, m_8 = 3$	8
3	30	$\{0.2; ...; 0.9\}$	$m_1, ..., m_5 = 4 \mid m_6, m_7 = 5$	7
4	40	$\{0.1; ...; 0.9\}$	$m_1, ..., m_{10} = 4$	10
5	50	$\{0.1; ...; 0.9\}$	$m_1, ..., m_{10} = 4 \mid m_{11}, m_{12} = 5$	12
6	60	$\{0.1; ...; 0.9\}$	$m_1, ..., m_{15} = 4$	15
7	70	$\{0.1\}$	$m_1, ..., m_{16} = 4 \mid m_{17}, m_{18} = 3$	18
8	70	$\{0.2; ...; 0.9\}$	$m_1, ..., m_{15} = 4 \mid m_{16}, m_{17} = 5$	17

Each line in Tables 4.2, 4.3 and 4.4 corresponds to a different design. This means that most of the designs for a particular sample size and scenario are optimal for various 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.

Table 4.2 shows that in the $(3, 3, 3, 3)$ and $(2, 3, 3, 4)$ scenarios, it is often statistically most efficient to administer three profiles to respondents. However, in the $(2, 3, 3, 4)$ scenario it is sometimes most efficient to administer four profiles to one or more respondents, especially when the correlation is increased. Also in the $(3, 3, 3, 3)$ scenario blocks of size four appear for higher correlations. The optimal number of respondents decreases in these cases. The result that sometimes more profiles are grouped when the correlation is increased is not unexpected as Goos and Vandebroek (2004) found a similar blocking pattern in the series of $\mathcal{D}$-optimal split-plot designs they computed.

While the optimal blocking structure of the $(2, 3, 3, 4)$ scenario corresponds to that of the $(3, 3, 3, 3)$ scenario, the optimal blocking structure of the $(2, 3, 3, 5)$ scenario largely differs from it. Table 4.3 shows that in general for the $(2, 3, 3, 5)$ scenario it is most efficient to assign sets of three and/or five profiles to the respondents. The lower the correlation, the more sets of size three are comprised. The higher the correlation, the more sets of size five are included. As a result, there is again a stronger grouping tendency at higher correlations, but it is more pronounced here than in the $(3, 3, 3, 3)$ and $(2, 3, 3, 4)$ scenarios. As can be seen, three or four design structures are possible for almost each sample size in the $(2, 3, 3, 5)$ scenario. This is more than the one or two design structures for each sample size in the $(3, 3, 3, 3)$ and $(2, 3, 3, 4)$ scenarios.

In the optimal blocking structure of the most heterogeneous level setting, the $(2, 3, 4, 5)$ scenario, contained in Table 4.4, the blocks of size three from in the previous scenarios mostly disappeared and were replaced by blocks of size four. It is thus generally most efficient to present four profiles to the respondents. Also, the stronger grouping tendency at higher correlations is

slightly apparent.

Note that the $\mathcal{D}$-optimal conjoint design for a given scenario, sample size and degree of correlation is not unique in its case. For each design problem a number of equivalent designs exist. This is because the $\mathcal{D}$-optimal conjoint designs are constructed for the random respondent effects model (4.2) embracing main effects only. As such, the $\mathcal{D}$-optimal conjoint design for 81 observations in the $(3, 3, 3, 3)$ scenario is equivalent to the 81-run $\mathcal{D}$-optimal 3^4 block design containing 27 blocks of size three. This design including all 81 candidate profiles appears in Table B1 of Appendix B. It is created using Appendix 5A on page 253 of Wu and Hamada (2000), but an equivalent design can also be generated using the interchange procedure with the 81 candidate profiles, block sizes of three and a value for $\rho \neq 0$ as an input. As another example, the $\mathcal{D}$-optimal conjoint design for 72 observations in the $(2, 3, 3, 4)$ scenario is equivalent to the $\mathcal{D}$-optimal arrangement of all 72 candidate profiles in blocks of size three. The interchange procedure yielded this arrangement which is shown in Table B2 of Appendix B.

4.6.4 Compromising between practical and optimal blocking structures

In this section, we investigate whether some of the optimal blocking structures 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 using the fixed block size algorithm given a more practical blocking structure as an input.

We tackle the following four cases each involving a different scenario. The first case concerns the conjoint designs in the $(3, 3, 3, 3)$ scenario 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 fewer respondent is needed. The second case covers the designs in the $(2, 3, 3, 4)$ scenario in which four profiles are presented to more than two respondents. It would be very convenient if the blocking structure of three profiles per respondent could be extended to these instances.

The third case involves all the designs in the $(2, 3, 3, 5)$ scenario as it would come in handy to have one fixed block size, a block size of either three or five, applicable to all examples. The fourth case considers the designs in the $(2, 3, 4, 5)$ scenario with blocks of size three. It would be very useful if the corresponding blocking structures could be replaced by the more frequent structures with blocks of size five.

We discuss these four cases more at length and investigate how much one loses in $\mathcal{D}$-efficiency by using $\mathcal{D}$-optimal conjoint designs with more practical blocking structures than the optimal ones. If the losses in $\mathcal{D}$-efficiency incurred are negligible, we retain the more practical structures. To determine the efficiency losses of using a $\mathcal{D}$-optimal conjoint design with a suboptimal blocking structure, we calculate how many observations would be saved if a $\mathcal{D}$-optimal conjoint design with an optimal blocking structure were applied whose $\mathcal{D}$-criterion value equals that of the suboptimal conjoint design. In other words, we express the losses in $\mathcal{D}$-efficiency in terms of the number of redundant observations of the $\mathcal{D}$-optimal conjoint design with the suboptimal blocking structure. In Appendix C we describe in detail how to compute the number of redundant observations.

Cases 1 & 2: General blocking structure in the $(3, 3, 3, 3)$ and $(2, 3, 3, 4)$ scenarios

In the $(3, 3, 3, 3)$ and $(2, 3, 3, 4)$ scenarios it is often most efficient to assign three profiles to the respondents. In some cases in the $(3, 3, 3, 3)$ scenario where the sample size is not a multiple of three, it is most efficient to administer one or two profiles to one of the respondents. This is the case for the designs with a sample size of 40 and 70 at lower correlations and for the designs with a sample size of 20 and 50. The assignment of one or two profiles to one respondent seems however not very attractive from a practical standpoint. Moreover, if there are costs associated with each respondent, it is also not very cost efficient. Therefore, we calculated how much one loses in $\mathcal{D}$-efficiency when assigning four profiles instead of three to one or two respondents so as to save on one respondent. It turns out that the losses in $\mathcal{D}$-efficiency from applying the more practical blocking structures are so small that there are no redundant observations. As a result, the structures

constitute a good compromise between practical and statistical efficiency.

Note that the assignment of four profiles to one or two respondents is most efficient for the remainder of the cases in the $(3, 3, 3, 3)$ scenario and for most of the cases in the $(2, 3, 3, 4)$ scenario where the sample size is not a multiple of three. An exception to the rule of administering three profiles to the respondents and four profiles to one or two respondents in case the sample size is not a multiple of three are the designs in the $(2, 3, 3, 4)$ scenario with a sample size of 30 and 40 at higher correlations. For these instances, it is most efficient to present four profiles to more than two respondents. Hence, we examined whether we can spread the profiles more equally over respondents in blocks of three. In the case of 40 observations this means that one respondent receives four profiles. Also here, the resulting losses in $\mathcal{D}$-efficiency turn out to be negligible as there are no redundant observations.

So in general, to construct conjoint designs for the $(3, 3, 3, 3)$ and $(2, 3, 3, 4)$ scenarios, it is efficient to use blocks of three profiles and one or two blocks of four profiles in case the sample size is not a multiple of three. This general blocking structure can be given as an input to the fixed block size algorithm to generate the $\mathcal{D}$-optimal conjoint designs. We found that this strategy is appropriate for the designs with a sample size larger than 10. Because the blocking structure is provided in the fixed block size algorithm, this algorithm takes much less computing time than the conjoint design algorithm (see also Section 4.6.5).

Case 3: General blocking structure in the $(2, 3, 3, 5)$ scenario

In the $(2, 3, 3, 5)$ scenario it is generally most efficient to group profiles in blocks of three and/or five. Instead of combining these two block sizes in function of the degree of correlation, it would be simpler to rely on one fixed block size, a block size of either three or five. Therefore, we examined whether it is feasible to use blocks of size three and one or two blocks of size five in case the sample size is not divisible by three. Also, we studied whether five profiles can be administered to the respondents and three profiles to a maximum of four respondents in case the sample size is not divisible by five. This latter blocking structure turns out to be the best as there are

no redundant observations in this case. On the other hand, the assignment of three profiles to the respondents results in some redundant observations at higher correlations. Consequently, since the losses in $\mathcal{D}$-efficiency are negligible from adopting a fixed block size of five, we can exploit this general block size in the fixed block size algorithm to construct $\mathcal{D}$-optimal conjoint designs for the $(2, 3, 3, 5)$ scenario. We found that this approach works well for the designs with a sample size larger than 11.

Case 4: General blocking structure in the $(2, 3, 4, 5)$ scenario

In the $(2, 3, 4, 5)$ scenario it is most efficient to spread the profiles over respondents in blocks of four. In most cases where the sample size is not a multiple of four, five profiles are assigned to a maximum of three respondents. We obtained this result by computing some additional $\mathcal{D}$-optimal conjoint designs for other sample sizes than the ones in Table 4.4. Sometimes, however, three profiles are administered to one or two respondents. This is the case for the designs with a sample size of 30 and 70 for a degree of correlation of 0.1. For these instances, we found that the use of blocks of size four and two blocks of size five yields no redundant observations. In general, the assignment of four profiles to the respondents and five profiles to a maximum of three respondents in case the sample size is not a multiple of four is efficient in all instances. As a result, to generate $\mathcal{D}$-optimal conjoint designs for the $(2, 3, 4, 5)$ scenario we can apply the fixed block size algorithm with the general blocking structure as an input. Note that only the designs with a sample size larger than 12 can be constructed in this way.

To conclude, for each of the four scenarios we found a general blocking structure supported by one predominant block size. As we show in the next section, the production of $\mathcal{D}$-optimal conjoint designs for fairly large sample sizes with the conjoint design algorithm takes a long time. Therefore, to construct large designs, we recommend to first compute some smaller $\mathcal{D}$-optimal conjoint designs with the conjoint design algorithm to derive the general blocking structure. This structure can then be provided to the fixed block size algorithm to compute the larger designs.

4.6.5 Computing times of $\mathcal{D}$-optimal conjoint designs

We now illustrate the huge time savings from applying the fixed block size algorithm with an efficient blocking structure as an input to generate $\mathcal{D}$-optimal conjoint designs for fairly large sample sizes. To do so, we compare the computing times to construct the $\mathcal{D}$-optimal conjoint designs in the $(3, 3, 3, 3)$ and $(2, 3, 3, 4)$ scenarios with the conjoint design algorithm and the fixed block size algorithm. We included the general blocking structure of three blocks per respondent in the fixed block size algorithm. We implemented both algorithms in Fortran 77 and registered the computing times using a Dell personal computer with a 1.80 GHz Intel Processor and 256 MB RAM. Table 4.5 shows the times per $1,000$ tries together with the numbers of tries used to create the $\mathcal{D}$-optimal conjoint designs for any degree of correlation ($\rho \neq 0$). We believe these numbers are sufficient so as not to be stuck in locally optimal designs.

Figure 4.2 contains the plot with the computing times per $1,000$ tries for the $\mathcal{D}$-optimal conjoint designs in the $(3, 3, 3, 3)$ scenario. We omitted the plot for the $(2, 3, 3, 4)$ scenario because it exhibits a similar pattern. Table 4.5 reveals that many more tries are required when the conjoint design algorithm is used instead of the fixed block size algorithm. In addition to that, Figure 4.2 shows that the computing times per $1,000$ tries with the conjoint design algorithm are long and grow exponentially with the sample size. In contrast, the computing times per $1,000$ tries with the fixed block size algorithm are much shorter and hardly increase with the sample size. As a result, it takes much more time to generate the $\mathcal{D}$-optimal conjoint designs with the conjoint design algorithm than with the fixed block size algorithm. Inversely put, the fixed block size algorithm dramatically speeds up the computing times. Particularly for the designs with a sample size of 40 and more enormous time savings are measured.

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

As an alternative way to quickly generate a relatively large $\mathcal{D}$-optimal conjoint design, a researcher can consider replicating a smaller $\mathcal{D}$-optimal conjoint design. If this design approach results in only negligible losses in $\mathcal{D}$-

Table 4.5: Computing times per $1,000$ tries and numbers of tries used to generate the $\mathcal{D}$-optimal conjoint designs in the $(3,3,3,3)$ and $(2,3,3,4)$ scenarios with the conjoint design algorithm and the fixed block size algorithm. The times are expressed in hours:minutes.

		Conjoint design algorithm		Fixed block size algorithm	
n	Scenario	time/$1,000$ tries	# tries	time/$1,000$ tries	# tries
20	$(3,3,3,3)$	01:17	$2,000$	00:02	$1,000$
20	$(2,3,3,4)$	01:05	$2,000$	00:01	$1,000$
24	$(3,3,3,3)$	02:04	$3,000$	00:02	$2,000$
24	$(2,3,3,4)$	01:24	$3,000$	00:02	$2,000$
30	$(3,3,3,3)$	02:54	$3,000$	00:03	$2,000$
30	$(2,3,3,4)$	02:02	$3,000$	00:03	$2,000$
36	$(3,3,3,3)$	03:58	$4,000$	00:04	$2,000$
36	$(2,3,3,4)$	02:47	$4,000$	00:03	$2,000$
40	$(3,3,3,3)$	04:52	$4,000$	00:05	$2,000$
40	$(2,3,3,4)$	03:05	$4,000$	00:04	$2,000$
50	$(3,3,3,3)$	08:31	$5,000$	00:08	$3,000$
50	$(2,3,3,4)$	04:36	$5,000$	00:07	$3,000$
60	$(3,3,3,3)$	11:05	$6,000$	00:10	$3,000$
60	$(2,3,3,4)$	07:14	$6,000$	00:09	$3,000$
70	$(3,3,3,3)$	15:42	$7,000$	00:15	$4,000$
70	$(2,3,3,4)$	09:18	$7,000$	00:15	$4,000$
72	$(3,3,3,3)$	16:44	$7,000$	00:15	$4,000$
72	$(2,3,3,4)$	10:43	$7,000$	00:15	$4,000$
81	$(3,3,3,3)$	21:19	$8,000$	00:21	$5,000$
81	$(2,3,3,4)$	12:51	$8,000$	00:18	$5,000$

efficiency, it would be quite desirable from a practical standpoint. This is because replicating a small $\mathcal{D}$-optimal conjoint design is cheaper both computationally and financially. The financial benefit is due to the fact that each distinct set of profiles can be rated by at least two respondents so that fewer profiles need to be manufactured. To evaluate the statistical efficiency of a design plan that contains one or more replications of a small $\mathcal{D}$-optimal con-

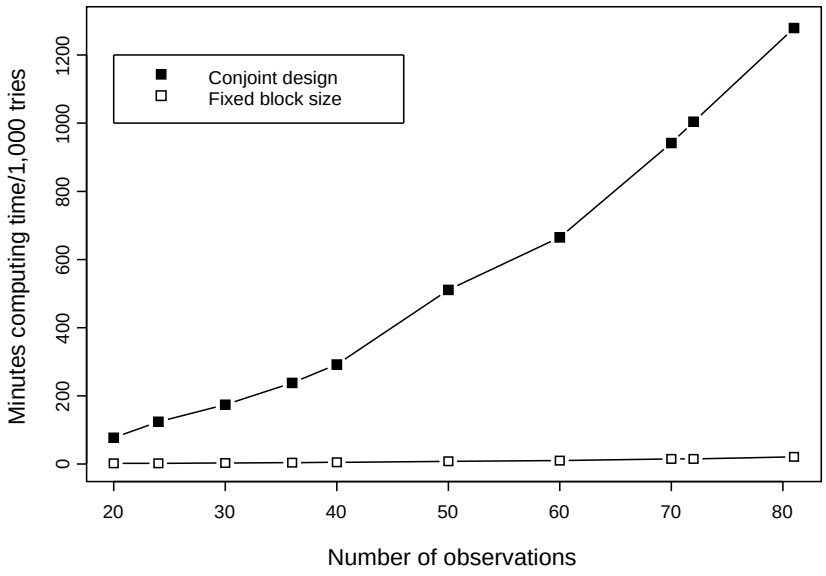


Figure 4.2: Minutes of computing time per $1,000$ tries to generate the $\mathcal{D}$-optimal conjoint designs in the $(3, 3, 3, 3)$ scenario using the conjoint design algorithm and the fixed block size algorithm.

joint design, we calculate the number of redundant observations associated with this plan. So we determine how many observations would be saved if a $\mathcal{D}$-optimal conjoint design were applied whose $\mathcal{D}$-criterion value equals that of the replicated design plan.

To derive the number of redundant observations of a replicated design plan, we need to compute the $\mathcal{D}$-criterion value of the replicated design plan. Evidently, we suspect all design plans studied in this section to be suboptimal. The $\mathcal{D}$-criterion value, $\mathcal{D}_n^{sub}$, of a design plan of n observations consisting of a multiple c of a small $\mathcal{D}$-optimal conjoint design with sample size n_s^* is given by

$$\mathcal{D}_n^{sub} = c \times \mathcal{D}_{n_s^*}^{opt}, \tag{4.24}$$

where $\mathcal{D}_{n_s^*}^{opt}$ denotes the $\mathcal{D}$-criterion value of the small $\mathcal{D}$-optimal conjoint design. Recall that the $\mathcal{D}$-criterion values of the $\mathcal{D}$-optimal conjoint designs in the $(3, 3, 3, 3)$, $(2, 3, 3, 4)$, $(2, 3, 3, 5)$ and $(2, 3, 4, 5)$ scenarios appear

in Table A1 of Appendix A.

Table 4.6 shows the replication schemes we set up for each of the four scenarios. For the $(3, 3, 3, 3)$ and $(2, 3, 3, 4)$ scenarios we investigated whether we could replicate the $\mathcal{D}$-optimal conjoint designs with a sample size of 20, 24, 30 and 36 to carry out experiments with 40, 60 and 72 observations. For the $(2, 3, 3, 5)$ and $(2, 3, 4, 5)$ scenarios we studied replications of the $\mathcal{D}$-optimal conjoint designs with a sample size of 20 and 30 to conduct experiments with 40 and 60 observations.

Table 4.6: Replication schemes of small $\mathcal{D}$-optimal conjoint designs in the $(3, 3, 3, 3)$, $(2, 3, 3, 4)$, $(2, 3, 3, 5)$ and $(2, 3, 4, 5)$ scenarios.

	n		
Scenario	40	60	72
$(3, 3, 3, 3)$ & $(2, 3, 3, 4)$	2×20	3×20 2×30	3×24 2×36
$(2, 3, 3, 5)$ & $(2, 3, 4, 5)$	2×20	3×20 2×30	

For the $(3, 3, 3, 3)$ and $(2, 3, 3, 4)$ scenarios it turns out that there are almost no redundant observations when the $\mathcal{D}$-optimal conjoint designs with a sample size of 20, 24, 30 and 36 are replicated. However, caution should be exercised in replicating these designs more than thrice as we observed one redundant observation from replicating the designs with 20 and 24 observations thrice. The more replications are made, the more one loses in $\mathcal{D}$-efficiency.

We observed similar results for the replicated design plans in the $(2, 3, 3, 5)$ and $(2, 3, 4, 5)$ scenarios. For most design plans consisting of a $\mathcal{D}$-optimal conjoint design with a sample size of 20 or 30 there are no redundant observations. An exception however are the design plans from the triple replication of the three $\mathcal{D}$-optimal conjoint designs with a sample size of 20 in the $(2, 3, 3, 5)$ scenario (see Table 4.3). These design plans have two redundant observations. This can be explained by the occurrence of blocks of size four in the optimal blocking structures of the $\mathcal{D}$-optimal conjoint designs at lower

correlations. In addition, these designs contain at most 20 different design profiles which is rather small.

We can conclude from these examples that it is efficient to replicate small $\mathcal{D}$-optimal conjoint designs for larger experiments if two conditions are met. First, the blocking structure of the small $\mathcal{D}$-optimal conjoint design matches the predominant blocking structure of the given scenario. Second, the small conjoint design contains an acceptable number of different profiles. The better this second condition is fulfilled, or the larger the sample size of the "small" design, the more replications can be made.

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

In practice, conjoint designs have often been constructed by generating a $\mathcal{D}$-optimal CRD and assigning the profiles at random to the respondents. Although this approach is very fast, we illustrate in this section that it is statistically inefficient. More specifically, we examine the performance of the $\mathcal{D}$-optimal CRDs in the $(3, 3, 3, 3)$ scenario when the profiles are randomly spread over respondents in blocks of three. In other words, we analyze each of the $\mathcal{D}$-optimal CRDs with the random respondent effects model (4.2) using blocks of three profiles and one or two blocks of four profiles in case the sample size is not a multiple of three.

To evaluate the $\mathcal{D}$-optimal CRDs, we randomly generated for each CRD $1,000$ profile arrangements according to the general blocking structure of three profiles per respondent. We then computed the $\mathcal{D}$-criterion values of these profile arrangements for each degree of correlation ($\rho \neq 0$). For each array of $1,000$ profile arrangements, we found that the $\mathcal{D}$-criterion values exhibit a bell-shaped pattern. Therefore, we compared the average $\mathcal{D}$-criterion values with the $\mathcal{D}$-criterion values of the $\mathcal{D}$-optimal conjoint designs that are listed in Table A1 of Appendix A and that are also plotted in Figure 4.1. The resulting average $\mathcal{D}$-efficiencies for each array of $1,000$ profile arrangements appear in Figure 4.3. They decrease with the degree of correlation from about 95% to about 65%, suggesting that randomly distributing CRD-

profiles over respondents is statistically not very efficient.

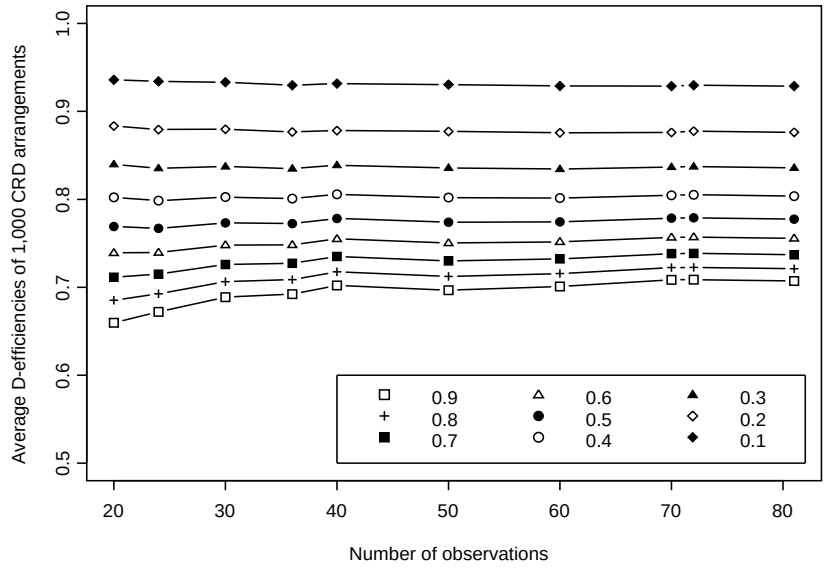


Figure 4.3: Average $\mathcal{D}$-efficiencies of $1,000$ profile arrangements of the $\mathcal{D}$-optimal CRDs in the $(3,3,3,3)$ scenario. Profiles are arranged according to the general blocking structure of three profiles per respondent.

To express the efficiency losses for each array of $1,000$ profile arrangements, we computed the numbers of redundant observations using the average $\mathcal{D}$-criterion values. The subsequent average numbers of redundant observations appear in Figure 4.4. As can be seen, these numbers 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 numbers of redundant observations. The minimum $\mathcal{D}$-criterion values result in the largest numbers of redundant observations whereas the maximum $\mathcal{D}$-criterion values result in the smallest numbers. The maxima and minima of redundant observations for degrees of correlation of 0.1 and 0.9 appear in Figure 4.5. They serve as bounds between which the maxima and minima of redundant observations for the other degrees of correlation are comprised. We observe that the max-

ima are more dispersed than the minima and that the spread between the maxima and minima increases with the degree of correlation.

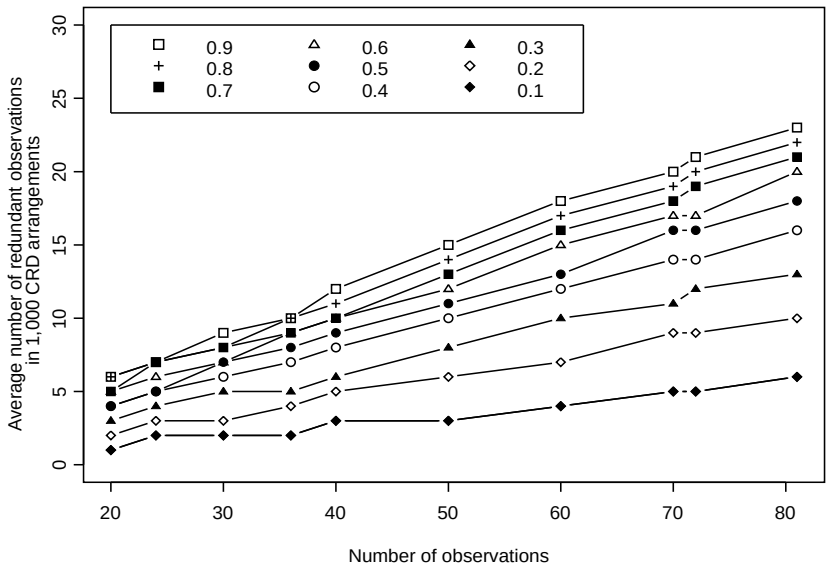


Figure 4.4: Average numbers of redundant observations of $1,000$ profile arrangements of the $\mathcal{D}$-optimal CRDs in the $(3,3,3,3)$ scenario. Profiles are arranged according to the general blocking structure of three profiles per respondent.

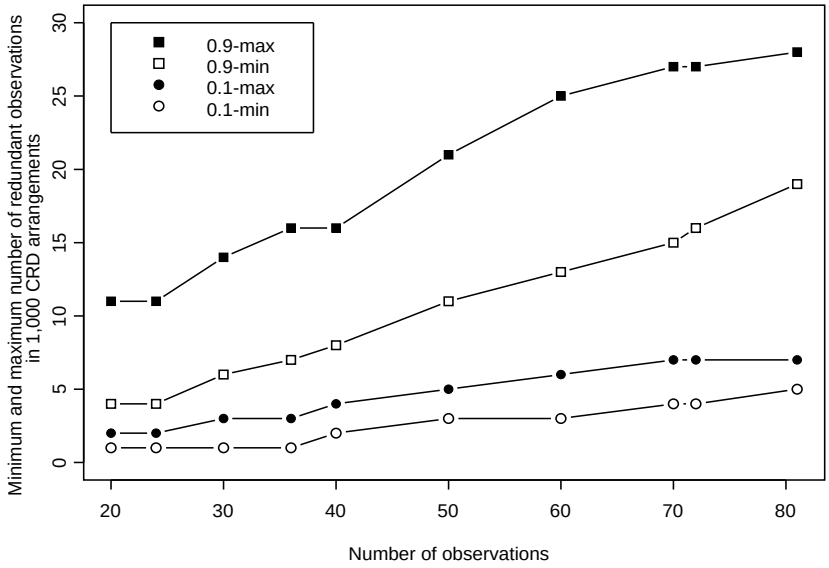


Figure 4.5: Minima and maxima of redundant observations of $1,000$ profile arrangements of the $\mathcal{D}$-optimal CRDs in the $(3, 3, 3, 3)$ scenario. Profiles are arranged according to the general blocking structure of three profiles per respondent for degrees of correlation of 0.1 and 0.9.

4.7 Conclusion

In this chapter, we constructed $\mathcal{D}$-optimal designs for conjoint experiments in which each respondent rates a small set of prototypes that must be tested in advance. Manufacturing these prototypes is expensive so that the number of design profiles is determined by the available budget. We used the linear model with a random effect to represent the fact that profile ratings from the same respondent are correlated. The resulting $\mathcal{D}$-optimal conjoint designs indicate how many respondents are necessary for a specific conjoint setting and which and how many profiles to administer to each of them.

We examined $\mathcal{D}$-optimal conjoint designs of various sample sizes at four level settings or scenarios. For each scenario, we were able to find a generally efficient blocking structure with which relatively large $\mathcal{D}$-optimal conjoint designs can be quickly constructed. Therefore, to generate large $\mathcal{D}$-optimal conjoint designs for any scenario, we recommend to first derive the general blocking structure for that scenario by constructing some smaller $\mathcal{D}$-optimal conjoint designs. The general blocking structure can then be given as an input to the design construction algorithm to produce the large designs. We conjecture that good sample sizes for conjoint designs for a given scenario are multiples of the least common multiple of the numbers of attribute levels. This is because the blocking structures of such designs seem to correspond to the general blocking structure of the scenario which we exploit in the design construction algorithm. Verifying or disproving this conjecture by a rigorous study is a potential future research topic.

Another way to compute larger $\mathcal{D}$-optimal conjoint designs in a time-efficient manner is to replicate a smaller $\mathcal{D}$-optimal design that has the general blocking structure of the scenario and a reasonable number of different profiles. This approach also requires fewer profiles to be manufactured because each distinct set of profiles can be rated by at least two respondents. This makes the experimental design cheaper financially. Finally, we demonstrated that constructing $\mathcal{D}$-optimal completely randomized designs and arbitrarily distributing the profiles to the respondents is statistically inefficient on average.

Appendix A. $\mathcal{D}$-criterion values of the $\mathcal{D}$-optimal conjoint designs and $\mathcal{D}$-optimal CRDs

Table A1: $\mathcal{D}$-criterion values of the $\mathcal{D}$-optimal conjoint designs ($\rho \neq 0$) and $\mathcal{D}$-optimal CRDs ($\rho = 0$) for the four scenarios described in Section 4.6.1.

	ρ									
n	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
a) $(3,3,3,3)$ scenario										
20	12.088	12.957	14.112	15.628	17.661	20.497	24.699	31.553	44.799	82.200
24	14.506	15.537	16.959	18.815	21.297	24.753	29.865	38.196	54.285	99.699
30	18.253	19.578	21.369	23.708	26.836	31.190	37.632	48.130	68.404	125.628
36	22.093	23.775	25.951	28.791	32.590	37.878	45.701	58.450	83.071	152.564
40	24.428	26.187	28.535	31.611	35.755	41.532	50.087	64.036	90.982	167.051
50	30.583	32.833	35.810	39.701	44.913	52.174	62.921	80.442	114.286	209.826
60	36.737	39.497	43.112	47.830	54.141	62.925	75.922	97.101	138.002	253.450
70	42.895	46.108	50.277	55.728	63.052	73.256	88.363	112.988	160.551	294.818
72	44.185	47.514	51.862	57.538	65.129	75.697	91.331	116.809	166.013	304.893
81	49.709	53.494	58.390	64.780	73.327	85.225	102.827	131.512	186.909	343.270
b) $(2,3,3,4)$ scenario										
20	11.247	11.952	12.937	14.263	16.068	18.603	22.376	28.543	40.478	74.200
24	13.768	14.632	15.826	17.435	19.627	22.710	27.299	34.807	49.338	90.407
30	17.088	18.158	19.637	21.632	24.349	28.171	33.862	43.189	61.249	112.283
36	20.696	22.007	23.810	26.237	29.538	34.181	41.090	52.392	74.266	136.086
40	22.908	24.315	26.297	28.973	32.617	37.762	45.414	57.927	82.141	150.566
50	28.709	30.500	32.986	36.345	40.919	47.353	56.932	72.598	102.920	188.612
60	34.505	36.682	39.679	43.716	49.212	56.940	68.448	87.273	123.709	226.684
70	40.238	42.758	46.243	50.947	57.354	66.365	79.781	101.726	144.200	264.238
72	41.449	44.051	47.643	52.489	59.087	68.368	82.186	104.787	148.534	272.173
81	46.577	49.503	53.542	58.987	66.403	76.832	92.361	117.761	166.925	305.873
c) $(2,3,3,5)$ scenario										
20	9.628	10.194	11.008	12.132	13.679	15.864	19.131	24.498	34.920	64.549
30	14.830	15.696	16.939	18.668	21.040	24.403	29.431	37.678	53.698	99.246
40	19.673	20.818	22.478	24.766	27.923	32.380	39.039	49.970	71.200	131.591
50	24.657	26.080	28.166	31.035	34.992	40.586	48.938	62.638	89.262	164.983
60	29.704	31.439	33.929	37.392	42.150	48.889	58.953	75.460	107.527	198.705
70	34.590	36.603	39.512	43.543	49.092	56.943	68.672	87.908	125.274	231.515
d) $(2,3,4,5)$ scenario										
20	8.867	9.444	10.256	11.355	12.853	14.961	18.106	23.268	33.319	62.027
30	13.413	14.244	15.436	17.075	19.318	22.478	27.196	34.943	50.027	93.119
40	18.057	19.208	20.844	23.066	26.099	30.374	36.754	47.228	67.601	125.875
50	22.532	23.949	25.980	28.745	32.523	37.847	45.792	58.839	84.242	156.810
60	27.196	28.917	31.371	34.709	39.268	45.692	55.281	71.026	101.685	189.270
70	31.639	33.633	36.484	40.368	45.672	53.147	64.303	82.621	118.289	220.184

Appendix B. $\mathcal{D}$-optimal conjoint designs for all candidate profiles

Table B1: $\mathcal{D}$-optimal conjoint design for all 81 candidate profiles in the $(3, 3, 3, 3)$ scenario.

	Attributes					Attributes					Attributes			
Set	1	2	3	4	Set	1	2	3	4	Set	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 B2: $\mathcal{D}$-optimal conjoint design for all 72 candidate profiles in the $(2, 3, 3, 4)$ scenario.

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

Appendix C. Number of redundant observations

The approach we propose to express losses in $\mathcal{D}$-efficiency when using suboptimal conjoint designs is a practical one and proceeds as follows. For each setting of the attribute levels and degree of correlation ($\rho \neq 0$), we perform a regression analysis of the $\mathcal{D}$-criterion values of the $\mathcal{D}$-optimal conjoint designs in Table A1 with respect to the sample size. We denote the sample size by n^* and the $\mathcal{D}$-criterion values by $\mathcal{D}_{n^*}^{opt}$. Each regression analysis yields an intercept and slope referred to as ψ and ω, respectively. Consider now a suboptimal conjoint design with sample size n and $\mathcal{D}$-criterion value $\mathcal{D}_n^{sub}$ for a given level setting and degree of correlation. It is clear that the $\mathcal{D}$-criterion value $\mathcal{D}_n^{sub}$ is smaller than the corresponding $\mathcal{D}$-criterion value $\mathcal{D}_{n^*}^{opt}$ where $n^* = n$. For the $\mathcal{D}$-criterion value $\mathcal{D}_n^{sub}$ we derive how many observations n^* are required in the optimal case using the equations

$$\mathcal{D}_n^{sub} = \mathcal{D}_{n^*}^{opt} \tag{C1}$$

and

$$\mathcal{D}_{n^*}^{opt} = \psi + \omega n^*. \tag{C2}$$

To further illustrate these expressions, we have drawn a regression line in Figure C1. Obviously, the sample size n^* is smaller than or, due to rounding, equal to the sample size n of the suboptimal conjoint design. The difference in numbers of observations, $n - n^*$, gives a clear indication of the extent of the efficiency losses. The larger the difference, the higher the losses. We refer to this difference as the number of redundant observations because it specifies how many observations would be saved if a $\mathcal{D}$-optimal conjoint design were applied to reach the same level of $\mathcal{D}$-efficiency as obtained by the suboptimal design.

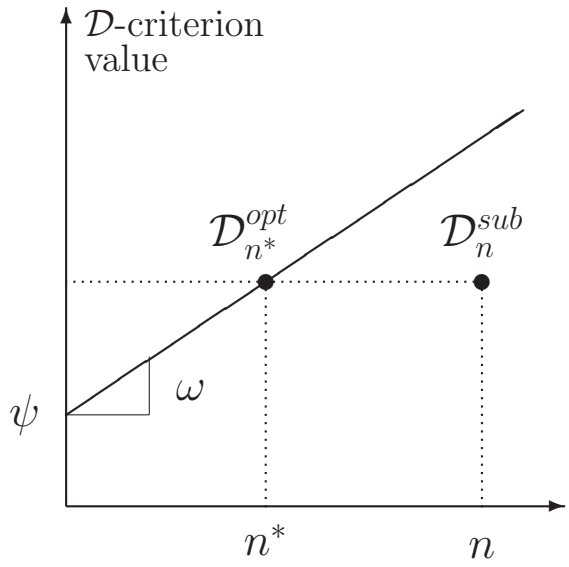


Figure C1: Derivation of the number of redundant observations, $n - n^*$, using regression analysis.

Chapter 5

Two-level variance-balanced rating-based conjoint designs

This chapter has been submitted as
☞ Kessels, R., Goos, P. and Vandebroek M. (2006). Optimal two-level conjoint designs for large numbers of attributes.

Abstract
In this chapter, we propose a simple strategy to construct $\mathcal{D}$-, $\mathcal{A}$-, $\mathcal{G}$- and $\mathcal{V}$-optimal two-level multi-attribute designs for rating-based conjoint studies. Our approach combines orthogonal designs and balanced or partially balanced incomplete block designs. In order not to overload respondents with complicated tasks, the designs hold one or more attributes at a constant level. The designs are variance-balanced meaning that they yield an equal amount of information on each of the part-worths. Some examples are provided to illustrate the method.

5.1 Introduction

A conjoint experiment tries to elucidate consumer preferences for the attributes of a good, that is a product or a service. This is usually done by asking respondents to rate a set of profiles or alternatives of the good. A profile involves a combination of levels of a set of predefined attributes of the good. Some of the administered profiles describe existing goods, whereas others describe hypothetical and possibly prospective goods. The objective of a conjoint experiment is to elicit as much information as possible on the utilities people derive from the attribute levels. These utilities are also called part-worths and correspond to the parameters of a statistical model. By means of accurate parameter estimates, precise predictions are aimed to depict consumers' purchasing behavior in a given market. Companies can then develop new goods that lead to a substantial rise in clientele.

To obtain precise parameter estimates, an efficient conjoint design needs to be constructed. The conjoint designs we set up in this chapter are intended for screening out the vital few important attributes from a group of many potential ones. The different attributes involved in the conjoint designs have two levels each. In most conjoint studies, see e.g. Danaher (1997), Pullman, Moore and Wardell (2002) and references therein, the conjoint design consists of blocks or sets of an equal number of profiles. Each of these sets is evaluated by a different respondent. To be able to estimate all part-worths, each respondent is administered a different set of profiles.

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

respondents, we confine ourselves to a maximum of ten constant attributes in the design.

Because each of the profile sets has a number of attributes at a constant level, the conjoint designs addressed in this chapter show some similarities to split-plot designs. Split-plot designs are heavily used in industry when the levels of some of the experimental factors are difficult or costly to change or control. These factors are called whole plot factors and are kept at a constant level for several observations in the design. The other factors whose levels may vary are called sub-plot factors. Analogous to our type of conjoint design is that a split-plot design consists of blocks or groups of runs with the whole plot factors acting at a constant level. These blocks are termed whole plots. A split-plot design differs from our conjoint design in that the whole plot factors are naturally the same in each whole plot, whereas the constant attributes in a conjoint design may vary from set to set. We refer to the work of Goos and Vandebroek (2001b; 2004) and Goos (2002, 2006a) for more details on split-plot designs.

Our approach to construct multi-attribute conjoint designs exploits the same linear model that is used for generating split-plot designs. In the split-plot setting, the model includes a random effect representing the whole plot variation. In our conjoint setting, the random effect is attributable to a respondent who rates a set of profiles. The reason is that respondents are assumed to be heterogeneous, meaning that profile ratings from the same respondent are more similar than profile ratings from different respondents. A random effect for each respondent accommodates this heterogeneity. Respondents are thereby supposed to be randomly selected from a prespecified market population. In the conjoint setting, we refer to the model as the linear random respondent effects model. Also Brazier, Roberts and Deverill (2002) and Kessels, Goos and Vandebroek (2004) adopted this model for conjoint applications. Because the conjoint designs in this chapter are set up for detecting the significant attributes from a large number of attributes, only main-effects models are considered.

Our design construction approach is conceptually easy to understand and generates designs with sets of 2, 4 and 8 profiles. In principle, our approach can produce designs with profile sets of sizes equal to a power of two. However, we restrict ourselves to a maximum size of eight as respondents can easily become fatigued by having to evaluate more than eight profiles. Related to conjoint designs with sets of two profiles are paired comparison designs (Grossmann, Holling and Schwabe 2002; Grossmann et al. 2005; Grasshoff et al. 2004; Street, Bunch and Moore 2001; Street and Burgess 2004). These designs also comprise sets of two profiles, but in contrast to conjoint designs, each respondent evaluates all the sets. This is done by specifying the preferred profile in each set, and possibly also the preference strength.

Street, Bunch and Moore (2001) demonstrated that holding the levels of one or more attributes constant in two-level paired comparison designs leads to information losses when main-effects models are considered. This result also applies to conjoint designs. Keeping the rating tasks manageable for the respondents thus comes at a loss of information on the part-worths. Since all part-worths are assumed to be on the same footing, we look for conjoint designs that spread the information losses evenly over each part-worth. In other words, we want to set up conjoint designs that provide an equal amount of information on each part-worth. These conjoint designs are called variance-balanced conjoint designs.

To allow for variance balance in the conjoint designs, the constant attributes have to differ between the profile sets in such a way that each attribute is constant in an equal number of sets. In case of more than one constant attribute, we obtain an appropriate pattern of constant attributes using a balanced or partially balanced incomplete block design (BIBD or PBIBD). The levels in these one-factor block designs, called treatments, each occur the same number of times, which paves the way for the production of variance-balanced conjoint designs. To ensure the optimality of the designs, meaning that they yield maximum information on each part-worth, we choose the design profiles so that they form an orthogonal array. These designs offer the advantage that parameter estimates are statistically independent of each other.

The combination of orthogonal designs and BIBDs or PBIBDs has also been employed by Green (1974) to develop multi-attribute choice experiments for an equal and unequal number of levels for the attributes. Yet, the resulting designs are not guaranteed to be optimal as Grossmann, Holling and Schwabe (2002) have shown. In choice experiments, respondents are administered a series of choice sets that each comprise several profiles. The respondents then indicate their preferred profile in every choice set. The underlying model is usually nonlinear, which requires the specification of prior parameter values before deriving the design. However, Green (1974) assumed zero prior values, hereby simplifying the nonlinear design problem to a linear one. Because of this assumption, the designs generated can also be utilized for conjoint experiments.

In one of his examples, Green (1974) constructed designs for partial profiles. These profiles are described by only a subset of the attributes. The levels of the other attributes are left unspecified. A BIBD is used to determine which of the attributes constitute the profiles. An orthogonal design is then assigned to each combination of attributes selected by the BIBD to form the profiles for these attributes. This approach yields variance-balanced designs when the attributes have two levels each. It best resembles our strategy since the attributes provided by the BIBD are the non-constant attributes in our conjoint setting. However, we also determine the levels of the other attributes, the constant attributes in our setting, so as to construct conjoint designs for full profiles. Comparing ratings from different profiles is impossible otherwise.

Based on a different approach, Street and Burgess (2004) generated optimal two-level paired comparison designs that can also be employed as variance-balanced conjoint designs for sets of two profiles. In this case, the paired comparison designs are set up for full profiles in which the levels of one or more attributes are held constant. They are referred to as constant difference pairs. Street and Burgess (2004) propose the use of generators and orthogonal designs to construct the pairs. They exploited the nonlinear Bradley-Terry model for choice experiments. Because zero prior parameter values are assumed, the model comes down to the linear model. The construction

of variance-balanced constant difference pairs is illustrated in an empirical study by Severin (2000).

The outline of the remainder of the chapter is as follows. Section 5.2 reviews the random respondent effects model used in conjoint design. In Section 5.3 we discuss the optimality conditions and in Section 5.4 we refine these conditions to deal with large numbers of attributes. We explain our design construction approach in Section 5.5 and describe the information content of the resulting designs in Section 5.6. Section 5.7 concludes the chapter and highlights some further research possibilities.

5.2 The random respondent effects model

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

$$U_{ij} = \mathbf{x}_{ij}'\boldsymbol{\beta} + \gamma_i + \varepsilon_{ij}. \tag{5.1}$$

In the model, $\mathbf{x}_{ij}$ is a $(k + 1) \times 1$ vector with a one as first element and the attribute levels describing profile j that is rated by respondent i as the remaining k elements. The attributes are quantitative or categorical factors that have two levels each. One level is coded as -1 and the other level as $+1$. Consequently, each of the individual attribute levels occupies one entry in the vector $\mathbf{x}_{ij}$. Moreover, only main-effects models are considered so that the number of attributes involved in the design amounts to k. $\boldsymbol{\beta} = [\beta_0, ..., \beta_k]'$ is the $(k+1) \times 1$ unknown parameter vector with β_0 the intercept and $\beta_1, ..., \beta_k$ the part-worths or weights attached to the attribute levels. γ_i represents the

random effect of respondent i and ε_{ij} is a random error term.

In matrix notation, model (5.1) becomes

$$\boldsymbol{U} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\varepsilon}, \tag{5.2}$$

where $\boldsymbol{U}$ is a vector of n profile ratings, 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 $n \times (k+1)$ and $n \times b$, respectively. $\mathbf{X}$ is given by $[\mathbf{X}_1', ..., \mathbf{X}_b']'$, where $\mathbf{X}_i = [\mathbf{x}_{i1}, ..., \mathbf{x}_{im}]'$ collects the m profiles rated by respondent i. $\mathbf{Z} = \mathbf{I}_b \otimes \mathbf{1}_m$, where $\otimes$ is the Kronecker product and $\mathbf{1}_m$ an $m \times 1$ vector of ones. It is assumed that

$$\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}_n \text{ and } \mathrm{Cov}(\boldsymbol{\varepsilon}) = \sigma^2_{\boldsymbol{\varepsilon}}\mathbf{I}_n, \tag{5.3}$$

$$\mathrm{E}(\boldsymbol{\gamma}) = \mathbf{0}_b \text{ and } \mathrm{Cov}(\boldsymbol{\gamma}) = \sigma^2_{\boldsymbol{\gamma}}\mathbf{I}_b, \tag{5.4}$$

$$\text{and } \mathrm{Cov}(\boldsymbol{\gamma}, \boldsymbol{\varepsilon}) = \mathbf{0}_{b\times n}, \tag{5.5}$$

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

$$\mathbf{V}_m = \sigma^2_{\boldsymbol{\varepsilon}}\mathbf{I}_m + \sigma^2_{\boldsymbol{\gamma}}\mathbf{1}_m\mathbf{1}_m' = \sigma^2_{\boldsymbol{\varepsilon}}\left(\mathbf{I}_m + \frac{\rho}{1-\rho}\mathbf{1}_m\mathbf{1}_m'\right). \tag{5.6}$$

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

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

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\,\boldsymbol{U}, \tag{5.7}$$

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

5.3 Design optimality

Our strategy to construct efficient multi-attribute conjoint designs is based on the information matrix $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$ for the random respondent effects model (5.2). Goos and Vandebroek (2001b) showed that the information matrix of a design $\mathbf{X}$ with profile sets of size m can be written as

$$\mathbf{X}'\mathbf{V}^{-1}\mathbf{X} = \sigma_{\boldsymbol{\varepsilon}}^{-2}\left\{\mathbf{X}'\mathbf{X} - \frac{\rho}{1+\rho(m-1)}\sum_{i=1}^{b}(\mathbf{X}_i'\mathbf{1}_m)(\mathbf{X}_i'\mathbf{1}_m)'\right\}. \tag{5.8}$$

For notational ease, we define the matrices $\mathbf{A}$ and $\mathbf{B}$ as

$$\mathbf{A} = \mathbf{X}'\mathbf{X} \quad \text{and} \quad \mathbf{B} = \sum_{i=1}^{b}(\mathbf{X}_i'\mathbf{1}_m)(\mathbf{X}_i'\mathbf{1}_m)', \tag{5.9}$$

so that the information matrix $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$ is of the form $\mathbf{A} - f(\rho)\mathbf{B}$.

The designs we derive are optimal with regard to four criteria that are functions of the information matrix (5.8) irrespective of the value of ρ. These criteria are the $\mathcal{D}$-, $\mathcal{A}$-, $\mathcal{G}$- and $\mathcal{V}$-optimality criteria. The $\mathcal{D}$- and $\mathcal{A}$-optimality criteria both are concerned with a precise estimation of the parameters $\boldsymbol{\beta}$ in model (5.2). A $\mathcal{D}$-optimal design minimizes the determinant of the variance-covariance matrix $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$ or, equivalently, maximizes the determinant of the information matrix (5.8). An $\mathcal{A}$-optimal design minimizes the trace of the variance-covariance matrix. The $\mathcal{G}$- and $\mathcal{V}$-optimality criteria are concerned with making precise response predictions. Because conjoint experiments particularly focus on producing precise predictions, these criteria are vital in design construction. The $\mathcal{G}$-optimality criterion seeks designs that minimize the maximum prediction variance over the region of interest, whereas the $\mathcal{V}$-optimality criterion seeks designs that minimize the average prediction variance over the region of interest.

To obtain conjoint designs that are $\mathcal{D}$-, $\mathcal{A}$-, $\mathcal{G}$- and $\mathcal{V}$-optimal, we construct the profile sets in such a way that the information matrix (5.8) is diagonal with elements that are as large as possible. The approach was motivated by Goos (2006b) who presented conditions for designing optimal two-level main-effect split-plot experiments. To find the conjoint design $\mathbf{X}$ corresponding

to a diagonal information matrix $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$ with the largest possible diagonal elements, both matrices $\mathbf{A}$ and $\mathbf{B}$ in (5.9) should be diagonal. The diagonal elements of $\mathbf{A}$ should be as large as possible, whereas those of $\mathbf{B}$ should be as small as possible.

In our construction method we present in Section 5.5, $\mathbf{X}$ will turn out to be an orthogonal array such as a full 2^k or fractional 2^{k-p} factorial design or a Plackett-Burman design. As a result, $\mathbf{A} = n\mathbf{I}_{k+1}$. Ideally, the profiles in the orthogonal design $\mathbf{X}$ are arranged in sets so that $\mathbf{B}$ is a zero matrix. However, because we are concerned with experiments with large numbers of attributes where some levels are held constant to simplify the respondents' task, this is impossible. Also, the first element of $\mathbf{B}$ corresponds to the intercept and equals bm^2 for every imaginable design. How best to assign these constant attribute levels is discussed in the next section.

5.4 Large numbers of attributes

In order not to overload the respondents with a heavy rating task, we hold one or more attributes at a fixed level in each of the profile sets. In other words, we introduce perfect level overlap for one or more attributes in each profile set. These attributes are the *constant attributes* and may differ from set to set. We denote the number of constant attributes by k_c. The remaining $k_v = k - k_c$ attributes, the levels of which may vary, are the *non-constant attributes*. The concepts of constant and non-constant attributes are illustrated with the conjoint design in Table 5.1. This design consists of $b = 12$ sets with $m = 2$ profiles each. It has $k = 6$ attributes, two of which are constant so that $k_c = 2$ and $k_v = 4$. The levels of the constant attributes are highlighted in grey. As can be seen, the constant and non-constant attributes may differ between the profile sets.

Table 5.1: Optimal variance-balanced conjoint design with $b = 12$ sets of $m = 2$ profiles, $k_c = 2$ constant attributes and $k_v = 4$ non-constant attributes.

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

Keeping the levels of one or more attributes fixed in the profile sets of a conjoint design reduces the amount of information that can be collected from the experiment. Also Street, Bunch and Moore (2001) observed that imposing perfect level overlap in two-level main-effect paired comparison designs leads to information losses with respect to the part-worths of the constant attributes in each of the profile sets. Without perfect level overlap constraints, most information is obtained when the levels of each attribute are maximum balanced, meaning that they occur with equal frequency in each profile set. Note that this is the opposite of perfect level overlap.

If no constant attributes were involved, maximum level balance would imply that $\mathbf{X}_i'\mathbf{1}_m = [m\ \mathbf{0}_{1\times k}]', i = 1, ..., b$, so that all elements of the matrix $\mathbf{B}$ in (5.9) are zero, except for the single element corresponding to the intercept in the upper left entry. This would maximize the information in the experiment. When k_c constant attributes are imposed on the design, $\mathbf{X}_i'\mathbf{1}_m = [m\ \mathbf{s}_i']', i = 1, ..., b$, where $\mathbf{s}_i$ is a $k \times 1$ vector containing at least k_c nonzero elements. As a result of that, $\mathbf{B}$ has additional nonzero elements that cause the loss of information. To minimize the number of nonzero elements in $\mathbf{B}$, the profiles in $\mathbf{X}$ forming an orthogonal array should be grouped such that maximum level balance is still preserved for the non-constant attributes in each profile set. In that case, $\mathbf{s}_i$ has exactly k_c nonzero elements. In the conjoint design of Table 5.1, the profiles represent a Plackett-Burman design and their grouping is characterized by maximum level balance for the non-constant attributes.

It is possible that the optimal designs do not spread the information losses from perfect level overlap evenly over all part-worths. Stated differently, the diagonal elements for the part-worths of the matrix $\mathbf{B}$ in (5.9), and thus of the information matrix (5.8), may not be the same. Still, we assume that all part-worths are equally important so that we look for optimal designs that yield the same amount of information on each part-worth. Each attribute should therefore be constant in an equal number of profile sets. Optimal designs that meet this requirement are called variance-balanced. The conjoint design of Table 5.1 is optimal and variance-balanced. As can be seen, each attribute acts as a constant attribute in four profile sets.

The fact that each attribute in the conjoint design of Table 5.1 is constant in an even number of profile sets is not surprising. This is generally the case for optimal variance-balanced conjoint designs. As we choose $\mathbf{X}$ to be an orthogonal array and arrange the profiles so that the levels of each non-constant attribute sum to zero in each profile set, the levels of each constant attribute necessarily sum to zero over the profile sets. This implies an even number of profile sets for each constant attribute to offset a -1 for a constant attribute in a profile set by a $+1$ for the same constant attribute in another profile set.

Now, with this framework in mind, we describe our strategy to set up optimal two-level variance-balanced conjoint designs in which a number of attributes are constant in each profile set. Compared with optimal designs without constant attributes, these designs give up some statistical efficiency to keep the rating tasks doable. Also, to show manageable profile sets, we disregard designs with more than $k_c = 10$ constant attributes or more than $k_v = 4$ non-constant attributes. To develop the optimal variance-balanced conjoint designs, we need to determine

1. the constant attributes in each profile set,

2. the levels of these constant attributes and

3. the levels of the non-constant attributes.

In the next section, we discuss our design construction approach taking into account this sequence of steps.

5.5 Design construction approach

In our design construction approach, we distinguish between $k_c = 1$ and $k_c > 1$ constant attributes. For both cases, we run through the three steps as listed above. We explain our strategy by some example designs that appear in Appendix A. Optimal variance-balanced conjoint designs for $k_c = 1$ and $k_c > 1$ appear in Appendix A.1 and Appendix A.2, respectively.

5.5.1 Optimal variance-balanced conjoint designs for $k_c = 1$

The optimal variance-balanced designs with $k_c = 1$ constant attribute in Appendix A.1 are the smallest ones that can be created for $k_v = 2, 3$ or 4 non-constant attributes and $m = 2, 4$ or 8 profiles per set. To construct these designs, we cover the three steps as follows.

Steps 1 & 2: Determining the constant attributes and their levels

When $k_c = 1$, determining the constant attributes and their levels is fairly straightforward. To allow for variance balance, each attribute should be constant in an equal number of profile sets. Also, each attribute should be constant in an even number of profile sets to have as many -1's as $+1$'s for their levels. In the optimal variance-balanced conjoint designs of Tables A1, A2, A4, A5, A7 and A8 each attribute is constant in two profile sets and in the designs of Tables A3 and A6 each attribute is constant in four profile sets.

Step 3: Determining the levels of the non-constant attributes

To determine the levels of the non-constant attributes, we draw on orthogonal subdesigns. For $k_v = 2, 3$ or 4 non-constant attributes and $m = 2, 4$ or 8 profiles per set, we constructed the smallest possible orthogonal subdesigns with an even number of profile sets. As such, we can match these subdesigns with the constant attributes to produce the conjoint designs of Appendix A.1. We illustrate the use of orthogonal subdesigns for each case.

The optimal variance-balanced conjoint designs in Tables A1 and A2 have $k_v = 2$ non-constant attributes and profile sets of sizes $m = 2$ and $m = 4$, respectively. The design in Table A1 is built on the orthogonal subdesign in Table 5.2a and the design in Table A2 is built on the orthogonal subdesign in Table 5.2b. The subdesign in Table 5.2a consists of two profile sets of size $m = 2$ and the subdesign in Table 5.2b consists of two profile sets of size $m = 4$. The former subdesign and each of the profile sets in the latter subdesign represent the full 2^2 factorial design.

Table 5.2: Orthogonal subdesigns with the levels of $k_v = 2$ non-constant attributes. The subdesigns consist of two profile sets each.

a) $m = 2$

	NC attr	
Set	1	2
1	-1	-1
1	$+1$	$+1$
2	-1	$+1$
2	$+1$	-1

b) $m = 4$

	NC attr	
Set	1	2
1	-1	-1
1	$+1$	$+1$
1	-1	$+1$
1	$+1$	-1
2	-1	-1
2	$+1$	$+1$
2	-1	$+1$
2	$+1$	-1

The subdesigns in Tables 5.2a and 5.2b are incorporated thrice in the conjoint designs of Tables A1 and A2, respectively. Since this is the required minimum to have each of the $k = 3$ attributes act as a constant attribute, the conjoint designs are the smallest ones that can be produced. They both involve six respondents, but can be replicated to set up larger conjoint designs in which the number of respondents is a multiple of six. The resulting designs are still optimal and variance-balanced. Note that we have not shown an optimal variance-balanced conjoint design with $k_v = 2$ non-constant attributes and sets of $m = 8$ profiles. This is because the profile sets in such design would contain only duplicated profiles which makes no sense in a conjoint study.

The optimal variance-balanced conjoint designs with $k_c = 1$ constant attribute in Tables A3, A4 and A5 have $k_v = 3$ non-constant attributes and profile sets of sizes $m = 2, 4$ and 8, respectively. The orthogonal subdesigns needed for their construction appear in Tables 5.3a, 5.3b and 5.3c. The subdesign in Table 5.3a embraces four profile sets of size $m = 2$ and the subdesigns in Tables 5.3b and 5.3c embrace two profile sets of sizes $m = 4$ and $m = 8$. The full 2^3 factorial design was exploited to build the subdesigns. It

is included once in the subdesigns of Tables 5.3a and 5.3b and twice in the subdesign of Table 5.3c.

Table 5.3: Orthogonal subdesigns with the levels of $k_v = 3$ non-constant attributes. Subdesign a) consists of four profile sets and subdesigns b) and c) consist of two profile sets.

a) $m = 2$

	NC attr		
Set	1	2	3
1	−1	−1	−1
1	+1	+1	+1
2	−1	−1	+1
2	+1	+1	−1
3	−1	+1	−1
3	+1	−1	+1
4	−1	+1	+1
4	+1	−1	−1

b) $m = 4$

	NC attr		
Set	1	2	3
1	−1	−1	−1
1	+1	+1	+1
1	−1	−1	+1
1	+1	+1	−1
2	−1	+1	−1
2	+1	−1	+1
2	−1	+1	+1
2	+1	−1	−1

c) $m = 8$

	NC attr		
Set	1	2	3
1	−1	−1	−1
1	+1	+1	+1
1	−1	−1	+1
1	+1	+1	−1
1	−1	+1	−1
1	+1	−1	+1
1	−1	+1	+1
1	+1	−1	−1
2	−1	−1	−1
2	+1	+1	+1
2	−1	−1	+1
2	+1	+1	−1
2	−1	+1	−1
2	+1	−1	+1
2	−1	+1	+1
2	+1	−1	−1

The subdesigns are incorporated four times in the conjoint designs of Tables A3, A4 and A5 to allow each of the $k = 4$ attributes to act as a constant attribute. As such, the conjoint designs are the smallest possible ones. For the design in Table A3 sixteen respondents are needed and for the designs in Tables A4 and A5 eight respondents are needed. To build larger optimal variance-balanced conjoint designs with $k_c = 1$ constant attribute and $k_v = 3$ non-constant attributes, replications of the design tables should be made.

The last optimal variance-balanced conjoint designs with $k_c = 1$ constant attribute are the ones in Tables A6, A7 and A8 containing $k_v = 4$ non-constant attributes and profile sets of sizes $m = 2$, 4 and 8, respectively. The designs are constructed using the orthogonal subdesigns listed in Tables 5.4a, 5.4b and 5.4c. Table 5.4a shows two possible subdesigns of four profile sets of size $m = 2$ and Table 5.4b shows two possible subdesigns of two profile sets of size $m = 4$. The subdesign in Table 5.4c has two profile sets of size $m = 8$. The combinations of the two possible subdesigns in Tables 5.4a and 5.4b and the subdesign in Table 5.4c represent the full 2^4 factorial design. The subdesigns in Tables 5.4a and 5.4b each constitute an orthogonal fraction of this design.

We produced the conjoint design in Table A6 by implementing the two possible subdesigns in Table 5.4a one after the other. In the same way, we constructed the conjoint design in Table A7 from the two subdesigns in Table 5.4b. Other replication structures of the subdesigns are also possible since the optimality of the designs is not affected by the choice of subdesigns used. The three conjoint designs are the smallest possible ones since five subdesigns are needed to allow each of the $k = 5$ attributes to act as a constant attribute. The design in Table A6 requires twenty respondents and the designs in Tables A7 and A8 require ten respondents. The designs can be replicated when larger conjoint experiments with $k_c = 1$ constant attribute and $k_v = 4$ non-constant attributes are desirable.

5.5.2 Optimal variance-balanced conjoint designs for $k_c > 1$

The optimal variance-balanced conjoint designs with $k_c > 1$ constant attributes in Appendix A.2 appear in the right panels of the tables. The other panels are meant to clarify our design construction approach. The conjoint design in the right panel of Table A9 is the same design as in Table 5.1 for $k_c = 2$ constant attributes, $k_v = 4$ non-constant attributes and $m = 2$ profiles per set. The next two designs in Tables A10 and A11 extend the starting example to $m = 4$ and $m = 8$ profiles per set, respectively. We first explain the main idea of our strategy using these three conjoint designs and then discuss the other conjoint designs of Appendix A.2.

Table 5.4: Orthogonal subdesigns with the levels of $k_v = 4$ non-constant attributes. Subdesigns i. and ii. of a) consist of four profile sets and subdesigns i. and ii. of b) and subdesign c) consist of two profile sets.

a) $m = 2$

		NC attr			
	Set	1	2	3	4
i.	1	−1	−1	−1	−1
	1	+1	+1	+1	+1
	2	−1	−1	+1	+1
	2	+1	+1	−1	−1
	3	−1	+1	−1	+1
	3	+1	−1	+1	−1
	4	−1	+1	+1	−1
	4	+1	−1	−1	+1
ii.	1	−1	−1	−1	+1
	1	+1	+1	+1	−1
	2	−1	−1	+1	−1
	2	+1	+1	−1	+1
	3	−1	+1	−1	−1
	3	+1	−1	+1	+1
	4	−1	+1	+1	+1
	4	+1	−1	−1	−1

b) $m = 4$

		NC attr			
	Set	1	2	3	4
i.	1	−1	−1	−1	−1
	1	+1	+1	+1	+1
	1	−1	−1	+1	+1
	1	+1	+1	−1	−1
	2	−1	+1	−1	+1
	2	+1	−1	+1	−1
	2	−1	+1	+1	−1
	2	+1	−1	−1	+1
ii.	1	−1	−1	−1	+1
	1	+1	+1	+1	−1
	1	−1	−1	+1	−1
	1	+1	+1	−1	+1
	2	−1	+1	−1	−1
	2	+1	−1	+1	+1
	2	−1	+1	+1	+1
	2	+1	−1	−1	−1

c) $m = 8$

	NC attr			
Set	1	2	3	4
1	−1	−1	−1	−1
1	+1	+1	+1	+1
1	−1	−1	+1	+1
1	+1	+1	−1	−1
1	−1	+1	−1	+1
1	+1	−1	+1	−1
1	−1	+1	+1	−1
1	+1	−1	−1	+1
2	−1	−1	−1	+1
2	+1	+1	+1	−1
2	−1	−1	+1	−1
2	+1	+1	−1	+1
2	−1	+1	−1	−1
2	+1	−1	+1	+1
2	−1	+1	+1	+1
2	+1	−1	−1	−1

Step 1: Determining the constant attributes

If $k_c > 1$, it may be quite cumbersome to look for combinations of constant attributes that lead to perfect level overlap of each attribute in an equal number of profile sets. This is particularly true when k_c is large and not a divisor of k. Therefore, we advocate the use of balanced or partially balanced incomplete block designs (BIBDs or PBIBDs) to determine patterns of constant attributes that result in variance-balanced conjoint designs. BIBDs and PBIBDs describe how to arrange the levels of a single qualitative factor, called treatments, in groups or blocks of a certain size. Each treatment thereby occurs an equal number of times in the entire design. In BIBDs, the number of times two different treatments occur together in a block is the same for all pairs of treatments. This is not true for PBIBDs which makes there are more and smaller PBIBDs than BIBDs for a given number of treatments and block size.

We refer to the work of Cochran and Cox (1957) and Cox (1958) for a general account of BIBDs and PBIBDs. The former authors depict some tables of BIBDs whereas the latter author describes a simple method to set up PBIBDs. More discussions and tables of BIBDs are provided by Abel and Greig (1996) and Mathon and Rosa (1996). PBIBDs are fully elaborated in the work of Shah and Sinha (1989) and Street and Street (1996), and catalogs of these designs can be found in the work of Clatworthy (1973) and on a website by Sinha (see the Bibliography).

In Tables 5.5a, 5.5b, 5.5c and 5.5d, three PBIBDs and one BIBD are listed for six treatments and block sizes of two. Each of these designs can be exploited to construct optimal variance-balanced conjoint designs with $k_c = 2$ constant attributes and $k_v = 4$ non-constant attributes. Each block in the PBIBDs and the BIBD determines which of the six attributes should be constant in a prespecified number of profile sets. So each block defines $k_c = 2$ constant attributes for a certain number of profile sets. Because each of the six treatments appears the same number of times in the PBIBDs and the BIBD, each of the $k = 6$ attributes is constant in an equal number of profile sets.

Table 5.5: a) PBIBD with 3 blocks, b) PBIBD with 9 blocks, c) PBIBD with 12 blocks and d) BIBD with 15 blocks. Each design has 6 treatments and block sizes of 2.

a) PBIBD

Block	Levels	
1	1	4
2	2	5
3	3	6

b) PBIBD

Block	Levels	
1	1	2
2	1	4
3	1	6
4	2	3
5	2	5
6	3	4
7	3	6
8	4	5
9	5	6

c) PBIBD

Block	Levels	
1	1	2
2	1	3
3	1	5
4	1	6
5	2	3
6	2	4
7	2	6
8	3	4
9	3	5
10	4	5
11	4	6
12	5	6

d) BIBD

Block	Levels	
1	1	2
2	1	3
3	1	4
4	1	5
5	1	6
6	2	3
7	2	4
8	2	5
9	2	6
10	3	4
11	3	5
12	3	6
13	4	5
14	4	6
15	5	6

The conjoint designs in Tables A9, A10 and A11 are built on the PBIBD of Table 5.5a. This is illustrated in the left panels of the tables. The PBIBD has three blocks of size two indicating three combinations of two constant attributes. The first block of the PBIBD tells us that Attributes 1 and 4 should be constant in some of the profile sets. This is done in Profile sets 1 to 4 for each of the conjoint designs. The second block determines Attributes 2 and 5 to be constant. These constant attributes are exploited in Profile sets 5 to 8. According to the last block, Attributes 3 and 6 should be constant. This is accomplished in Profile sets 9 to 12. So each block of constant attributes is embedded in four profile sets in each of the conjoint designs. How to decide on the number of profile sets in which the same attributes are constant is laid out in the next section.

Steps 2 & 3: Determining the levels of the constant and non-constant attributes

The optimal variance-balanced conjoint designs in Appendix A.2 are all developed from PBIBDs and are the smallest ones that can be produced. For the conjoint designs in Tables A9, A10 and A11, we showed that the constant attributes are dictated by the three blocks in the PBIBD of Table 5.5a. The question now is at what levels these attributes have to be set and how often they have to be held fixed. To answer this question, we use intermediate designs, so-called constant attribute designs that are orthogonal for the constant attributes. These constant attribute designs appear in the middle panel of each design table.

To construct the constant attribute designs, we exploit the full 2^{k_c} factorial design or an orthogonal fraction of the full 2^{k_c} factorial design. The sample size of the full 2^{k_c} factorial design or the orthogonal fraction then determines the number of profile sets for each combination of constant attributes. The constant attribute designs for the conjoint designs in Tables A9, A10 and A11 utilize the full 2^2 factorial design for each combination of constant attributes. This means that $2^2 = 4$ profile sets have the same attributes constant with levels that fit the full 2^2 factorial design.

The last problem to solve before the design construction is complete is the determination of the levels of the non-constant attributes. Therefore, we call on the orthogonal subdesigns for the k_v non-constant attributes given by the tables in Section 5.5.1. In the conjoint design of Table A9, the two possible subdesigns of Table 5.4a are implemented. Since three subdesigns can be accommodated, one subdesign is implemented once and the other twice. They nicely fit in the four profile sets from the full 2^2 factorial design corresponding to a combination of constant attributes. Similarly, in the conjoint design of Table A10 each of the subdesigns in Table 5.4b is implemented thrice and in the conjoint design of Table A11 the subdesign in Table 5.4c is implemented six times.

In some cases if $k_c = 2$, the two non-orthogonal fractions of the full 2^2 factorial design in Table 5.6 provide another way to determine the levels of the

constant attributes. These fractions require only two profile sets for each combination of constant attributes and thus allow for smaller conjoint designs when $k_c = 2$. They both have one attribute at a fixed level so that they need to be jointly implemented to offset the fixed level of -1 by the fixed level of $+1$. As a result, the fractions can only be employed if the number of blocks in the BIBD or PBIBD is even and the orthogonal subdesigns for the k_v non-constant attributes enclose two profile sets. The use of the fractions is illustrated in the conjoint designs of Tables A12 and A13 for $k_c = 2$ constant attributes, $k_v = 2$ non-constant attributes and $m = 2$ and $m = 4$ profiles per set, respectively. The PBIBD specifies four combinations of constant attributes, the levels of which are determined by the non-orthogonal fractions. The subdesigns in Tables 5.2a and 5.2b fill out the levels of the non-constant attributes.

Table 5.6: Non-orthogonal fractions of the full 2^2 factorial design for constructing constant attribute designs for $k_c = 2$ constant attributes.

		Cst attr	
	Set	1	2
i	1	-1	-1
	2	-1	$+1$
ii	1	$+1$	-1
	2	$+1$	$+1$

To further illustrate our design construction approach, we discuss the optimal variance-balanced conjoint designs in Tables A14 and A15. They both have $m = 2$ profiles per set. The design in Table A14 has $k_c = 3$ constant attributes and $k_v = 3$ non-constant attributes and the design in Table A15 has $k_c = 4$ constant attributes and $k_v = 4$ non-constant attributes. To fix the levels of the constant attributes, minimum-size orthogonal fractions of the full 2^{k_c} factorial design are incorporated. These fractions are the smallest possible ones that allow the estimation of all k_c main effects. If $3 \leq k_c \leq 10$, minimum-size orthogonal fractions for the k_c constant attributes can be constructed using the generators in Table 5.7. We selected these generators from

Appendix 4A on pages 193–194 of Wu and Hamada (2000). Other generators that yield larger orthogonal fractions can also be retrieved in this appendix.

Table 5.7: Generators for constructing minimum-size orthogonal fractions for k_c attributes.

k_c	Size	Generators
3	$2^{3-1} = 4$	$3 = 12$
4	$2^{4-1} = 8$	$4 = 123$
5	$2^{5-2} = 8$	$4 = 12$, $5 = 13$
6	$2^{6-3} = 8$	$4 = 12$, $5 = 13$, $6 = 23$
7	$2^{7-4} = 8$	$4 = 12$, $5 = 13$, $6 = 23$, $7 = 123$
8	$2^{8-4} = 16$	$5 = 123$, $6 = 124$, $7 = 134$, $8 = 234$
9	$2^{9-5} = 16$	$5 = 123$, $6 = 124$, $7 = 134$, $8 = 234$, $9 = 1234$
10	$2^{10-6} = 16$	$5 = 123$, $6 = 124$, $7 = 134$, $8 = 234$, $9 = 1234$, $10 = 34$

The orthogonal fraction has a size of 2^{k_c-p}, where 2^{-p} refers to the fraction of the full 2^{k_c} factorial design. The full 2^{k_c-p} factorial design is used to determine the levels of the first $k_c - p$ constant attributes. The levels of the remaining p constant attributes are specified by the generators. For example, in case $k_c = 3$, the generator is $3 = 12$ meaning that the level of the third constant attribute is obtained by multiplying the levels of the first and second constant attribute. In this way, one orthogonal fraction is constructed. Other orthogonal fractions of the full 2^{k_c} factorial design can be produced by using one or more generators with a minus sign. For $k_c = 3$, $3 = -12$ is the generator of the second and only remaining orthogonal fraction.

We derive from Table 5.7 that there are also two minimum-size orthogonal fractions for $k_c = 4$ constant attributes. In the constant attribute design of Table A14 each of two minimum-size orthogonal fractions for $k_c = 3$ constant attributes is used twice and in the constant attribute design of Table A15 each of two minimum-size orthogonal fractions for $k_c = 4$ constant attributes is used thrice. The orthogonal fractions for $k_c = 3$ involve $2^{3-1} = 4$ profile sets for each combination of constant attributes and the orthogonal fractions for

$k_c = 4$ involve $2^{4-1} = 8$ profile sets. The levels of the non-constant attributes in the conjoint designs are set by means of the subdesign in Table 5.3a for $k_v = 3$ and the two possible subdesigns in Table 5.4a for $k_v = 4$.

5.6 Information content

In this last section, we discuss the information content of optimal variance-balanced conjoint designs with a varying number of constant attributes. More specifically, we are interested in how much we lose in terms of information by having one attribute constant, two attributes constant, and so forth.

For our study, we examine the designs in Appendix A that have $k = 4$ attributes. These are the designs in Tables A3, A4 and A5 of Appendix A.1 for $k_c = 1$ constant attribute and $k_v = 3$ non-constant attributes and the designs in Tables A12 and A13 of Appendix A.2 for $k_c = 2$ constant attributes and $k_v = 2$ non-constant attributes. The designs with $k_c = 1$ constant attribute have profile sets of sizes $m = 2, 4$ and 8, respectively, and the designs with $k_c = 2$ constant attributes have profile sets of sizes $m = 2$ and $m = 4$. We compare the information content of these designs with that of the optimal conjoint designs in which no constant attributes are used. These latter conjoint designs are nothing but orthogonally blocked two-level designs. They can be constructed using the generators in Appendix 3A on pages 150–151 and Appendix 4B on pages 199–203 of Wu and Hamada (2000).

In general, the information matrix (5.8) for optimal variance-balanced conjoint designs is diagonal with maximal diagonal elements that are equal for the part-worths. In Appendix B, we computed the information matrix for the design in Table A12. Based on these computations, we derived the amount of information on the intercept and part-worths for the other conjoint designs. The results for all five conjoint designs appear in Table 5.8.

For a given design setting, the optimal conjoint design without constant attributes leads to the same amount of information on the intercept as provided by the optimal conjoint design with constant attributes. However, as we men-

tioned in Sections 5.3 and 5.4, the amount of information on each part-worth of the optimal designs without constant attributes is larger. Because the levels of each attribute are maximum balanced in each of the profile sets, the part-worth elements of the matrix $\mathbf{B}$ in (5.9) are zero. As a result, the amount of information on each part-worth of the optimal designs without constant attributes is equal to n, the number of design profiles. For each of the five design cases in Table 5.8, we compared the value of n with the values for the part-worths in the table for 10 degrees of correlation $\rho \in \{0; 0.1; ...; 0.9\}$. As such, we computed the percentage information losses for each part-worth from imposing constant attributes. We plotted them in Figure 5.1.

Table 5.8: Amount of information on the intercept and part-worths of the optimal variance-balanced conjoint designs for a) $k_c = 1$ and $k_v = 3$ contained in Tables A3, A4 and A5 and for b) $k_c = k_v = 2$ contained in Tables A12 and A13.

	Conjoint design				Amount of information on	
	Table	m	b	n	the intercept	each part-worth
a) $k_c = 1$	A3	2	16	32	$32(1-\rho)/(1+\rho)$	$8(1-\rho)/(1+\rho)+24$
$k_v = 3$	A4	4	8	32	$32(1-\rho)/(1+3\rho)$	$8(1-\rho)/(1+3\rho)+24$
	A5	8	8	64	$64(1-\rho)/(1+7\rho)$	$16(1-\rho)/(1+7\rho)+48$
b) $k_c = 2$	A12	2	8	16	$16(1-\rho)/(1+\rho)$	$8(1-\rho)/(1+\rho)+8$
$k_v = 2$	A13	4	8	32	$32(1-\rho)/(1+3\rho)$	$16(1-\rho)/(1+3\rho)+16$

A close look at Figure 5.1 reveals that, given $m = 2$ or $m = 4$ and a value of ρ, the information losses for each part-worth of the conjoint designs with $k_c = 2$ constant attributes are twice the information losses of the conjoint designs with $k_c = 1$ constant attribute. We verified with some additional computations that in general, if the number of attributes, k, is fixed in a series of optimal two-level variance-balanced conjoint designs, the information losses are proportional to the number of constant attributes, k_c, used.

Figure 5.1 also shows that, given k_c and ρ, the information losses for each part-worth increase with m. As the profile sets of the conjoint designs get larger or the number of respondents drops, the information losses from con-

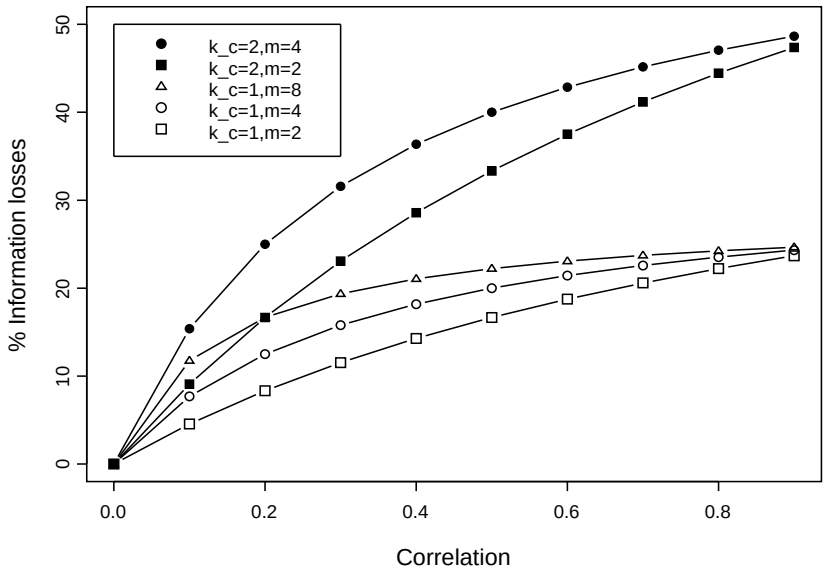


Figure 5.1: % Information losses for each part-worth from using $k_c = 1$ and $k_c = 2$ constant attributes for degrees of correlation ranging from 0 to 0.9. The five designs of Table 5.8 are considered.

stant attributes increase. This can also be concluded from $c_8 < c_4 < c_2 < 1$, where $c_m = (1 - \rho)/(1 + (m - 1)\rho)$ (see Appendix B). It implies that for a given number of design profiles, n, sets with $m = 2$ profiles are more efficient than sets with $m = 4$ profiles, which in their turn, are more efficient than sets with $m = 8$ profiles. For example, the designs in Tables A3 and A4 have equal sample sizes and equal numbers of profiles in which each attribute is constant. However, because the former design is constructed for $m = 2$ and the latter for $m = 4$, the information losses of the latter design are larger.

A last observation from Figure 5.1 is that, given m and k_c, the information losses for each part-worth increase with ρ. The more heterogeneous respondents are, the more information one loses by keeping the levels of one or more attributes constant. The increase of the losses with ρ is fairly linear if $m = 2$, but becomes quadratic for larger values of m. Note that if $\rho = 0$, respondents are assumed to be homogeneous so that the grouping of the profiles in sets, and thus also the matrix $\mathbf{B}$ in (5.9), does not matter anymore.

5.7 Conclusion

In this chapter, we presented a simple approach to construct optimal two-level conjoint designs that involve a large number of attributes. To reduce the cognitive burden on the respondents, the designs hold the levels of one or more attributes constant in each of the profile sets. The designs have sets of 2, 4 or 8 profiles and are optimal with respect to the $\mathcal{D}$-, $\mathcal{A}$-, $\mathcal{G}$- and $\mathcal{V}$-optimality criteria for estimating main-effects models.

The optimal conjoint designs have a diagonal information matrix collecting maximum information. In addition, the conjoint designs are variance-balanced meaning that they yield an equal amount of information on each of the part-worths. Also, the conjoint designs do not depend on the extent to which respondents are heterogeneous, as expressed by the degree of correlation. This makes them very practical to use. Prior to constructing the conjoint designs, we advise practitioners to think carefully about the number of constant attributes they want. Compared with optimal conjoint designs without constant attributes, optimal conjoint designs with constant attributes lead to information losses that are proportional to the number of constant attributes.

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

The design examples provided in Appendix A are the smallest ones supported by our approach. There are generally three ways to obtain larger conjoint designs. A first option is to choose a larger BIBD or PBIBD that defines more combinations of constant attributes. A second possibility is to employ larger orthogonal designs to fix the levels of the k_c constant attributes. Lastly, a

small conjoint design may be replicated.

As an additional justification of our method, it is interesting to note that the optimal variance-balanced conjoint designs with sets of two profiles are similar to the variance-balanced constant difference pairs constructed by Severin (2000). For example, the optimal paired comparison design with four attribute level differences in the work of Severin (2000, page 142) is equivalent to the duplicated conjoint design of our starting example in Table 5.1. In other words, Severin's (2000, page 142) design has the same structure as the conjoint design in Table 5.1 but contains twice as many profile sets.

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

Appendix A. Optimal variance-balanced conjoint designs

A.1. Designs with one constant attribute: $k_c = 1$

Table A1: Smallest optimal variance-balanced conjoint design with sets of $m = 2$ profiles, $k_c = 1$ constant attribute and $k_v = 2$ non-constant attributes.

	Attributes				Attributes				Attributes		
Set	1	2	3	Set	1	2	3	Set	1	2	3
1	−1	−1	−1	3	−1	−1	−1	5	−1	−1	−1
1	−1	+1	+1	3	+1	−1	+1	5	+1	+1	−1
2	+1	−1	+1	4	−1	+1	+1	6	−1	+1	+1
2	+1	+1	−1	4	+1	+1	−1	6	+1	−1	+1

Table A2: Smallest optimal variance-balanced conjoint design with sets of $m = 4$ profiles, $k_c = 1$ constant attribute and $k_v = 2$ non-constant attributes.

	Attributes				Attributes				Attributes		
Set	1	2	3	Set	1	2	3	Set	1	2	3
1	−1	−1	−1	3	−1	−1	−1	5	−1	−1	−1
1	−1	+1	+1	3	+1	−1	+1	5	+1	+1	−1
1	−1	−1	+1	3	−1	−1	+1	5	−1	+1	−1
1	−1	+1	−1	3	+1	−1	−1	5	+1	−1	−1
2	+1	−1	−1	4	−1	+1	−1	6	−1	−1	+1
2	+1	+1	+1	4	+1	+1	+1	6	+1	+1	+1
2	+1	−1	+1	4	−1	+1	+1	6	−1	+1	+1
2	+1	+1	−1	4	+1	+1	−1	6	+1	−1	+1

Table A3: Smallest optimal variance-balanced conjoint design with sets of $m = 2$ profiles, $k_c = 1$ constant attribute and $k_v = 3$ non-constant attributes.

	Attributes					Attributes			
Set	1	2	3	4	Set	1	2	3	4
1	−1	−1	−1	−1	9	−1	−1	−1	−1
1	−1	+1	+1	+1	9	+1	+1	−1	+1
2	−1	−1	−1	+1	10	−1	−1	−1	+1
2	−1	+1	+1	−1	10	+1	+1	−1	−1
3	+1	−1	+1	−1	11	−1	+1	+1	−1
3	+1	+1	−1	+1	11	+1	−1	+1	+1
4	+1	−1	+1	+1	12	−1	+1	+1	+1
4	+1	+1	−1	−1	12	+1	−1	+1	−1
5	−1	−1	−1	−1	13	−1	−1	−1	−1
5	+1	−1	+1	+1	13	+1	+1	+1	−1
6	−1	−1	−1	+1	14	−1	−1	+1	−1
6	+1	−1	+1	−1	14	+1	+1	−1	−1
7	−1	+1	+1	−1	15	−1	+1	−1	+1
7	+1	+1	−1	+1	15	+1	−1	+1	+1
8	−1	+1	+1	+1	16	−1	+1	+1	+1
8	+1	+1	−1	−1	16	+1	−1	−1	+1

Table A4: Smallest optimal variance-balanced conjoint design with sets of $m = 4$ profiles, $k_c = 1$ constant attribute and $k_v = 3$ non-constant attributes.

	Attributes					Attributes			
Set	1	2	3	4	Set	1	2	3	4
1	−1	−1	−1	−1	5	−1	−1	−1	−1
1	−1	+1	+1	+1	5	+1	+1	−1	+1
1	−1	−1	−1	+1	5	−1	−1	−1	+1
1	−1	+1	+1	−1	5	+1	+1	−1	−1
2	+1	−1	+1	−1	6	−1	+1	+1	−1
2	+1	+1	−1	+1	6	+1	−1	+1	+1
2	+1	−1	+1	+1	6	−1	+1	+1	+1
2	+1	+1	−1	−1	6	+1	−1	+1	−1
3	−1	−1	−1	−1	7	−1	−1	−1	−1
3	+1	−1	+1	+1	7	+1	+1	+1	−1
3	−1	−1	−1	+1	7	−1	−1	+1	−1
3	+1	−1	+1	−1	7	+1	+1	−1	−1
4	−1	+1	+1	−1	8	−1	+1	−1	+1
4	+1	+1	−1	+1	8	+1	−1	+1	+1
4	−1	+1	+1	+1	8	−1	+1	+1	+1
4	+1	+1	−1	−1	8	+1	−1	−1	+1

Table A5: Smallest optimal variance-balanced conjoint design with sets of $m = 8$ profiles, $k_c = 1$ constant attribute and $k_v = 3$ non-constant attributes.

	Attributes					Attributes			
Set	1	2	3	4	Set	1	2	3	4
1	−1	−1	−1	−1	5	−1	−1	−1	−1
1	−1	+1	+1	+1	5	+1	+1	−1	+1
1	−1	−1	−1	+1	5	−1	−1	−1	+1
1	−1	+1	+1	−1	5	+1	+1	−1	−1
1	−1	−1	+1	−1	5	−1	+1	−1	−1
1	−1	+1	−1	+1	5	+1	−1	−1	+1
1	−1	−1	+1	+1	5	−1	+1	−1	+1
1	−1	+1	−1	−1	5	+1	−1	−1	−1
2	+1	−1	−1	−1	6	−1	−1	+1	−1
2	+1	+1	+1	+1	6	+1	+1	+1	+1
2	+1	−1	−1	+1	6	−1	−1	+1	+1
2	+1	+1	+1	−1	6	+1	+1	+1	−1
2	+1	−1	+1	−1	6	−1	+1	+1	−1
2	+1	+1	−1	+1	6	+1	−1	+1	+1
2	+1	−1	+1	+1	6	−1	+1	+1	+1
2	+1	+1	−1	−1	6	+1	−1	+1	−1
3	−1	−1	−1	−1	7	−1	−1	−1	−1
3	+1	−1	+1	+1	7	+1	+1	+1	−1
3	−1	−1	−1	+1	7	−1	−1	+1	−1
3	+1	−1	+1	−1	7	+1	+1	−1	−1
3	−1	+1	+1	−1	7	−1	+1	−1	−1
3	+1	+1	−1	+1	7	+1	−1	+1	−1
3	−1	+1	+1	+1	7	−1	+1	+1	−1
3	+1	+1	−1	−1	7	+1	−1	−1	−1
4	−1	−1	−1	−1	8	−1	−1	−1	+1
4	+1	−1	+1	+1	8	+1	+1	+1	+1
4	−1	−1	−1	+1	8	−1	−1	+1	+1
4	+1	−1	+1	−1	8	+1	+1	−1	+1
4	−1	+1	+1	−1	8	−1	+1	−1	+1
4	+1	+1	−1	+1	8	+1	−1	+1	+1
4	−1	+1	+1	+1	8	−1	+1	+1	+1
4	+1	+1	−1	−1	8	+1	−1	−1	+1

Table A6: Smallest optimal variance-balanced conjoint design with sets of $m = 2$ profiles, $k_c = 1$ constant attribute and $k_v = 4$ non-constant attributes.

	Attributes						Attributes				
Set	1	2	3	4	5	Set	1	2	3	4	5
1	−1	−1	−1	−1	−1	11	−1	+1	+1	−1	+1
1	−1	+1	+1	+1	+1	11	+1	−1	+1	+1	−1
2	−1	−1	−1	+1	+1	12	−1	+1	+1	+1	−1
2	−1	+1	+1	−1	−1	12	+1	−1	+1	−1	+1
3	+1	−1	+1	−1	+1	13	−1	−1	−1	−1	+1
3	+1	+1	−1	+1	−1	13	+1	+1	+1	−1	−1
4	+1	−1	+1	+1	−1	14	−1	−1	+1	−1	−1
4	+1	+1	−1	−1	+1	14	+1	+1	−1	−1	+1
5	−1	−1	−1	−1	+1	15	−1	+1	−1	+1	−1
5	+1	−1	+1	+1	−1	15	+1	−1	+1	+1	+1
6	−1	−1	−1	+1	−1	16	−1	+1	+1	+1	+1
6	+1	−1	+1	−1	+1	16	+1	−1	−1	+1	−1
7	−1	+1	+1	−1	−1	17	−1	−1	−1	−1	−1
7	+1	+1	−1	+1	+1	17	+1	+1	+1	+1	−1
8	−1	+1	+1	+1	+1	18	−1	−1	+1	+1	−1
8	+1	+1	−1	−1	−1	18	+1	+1	−1	−1	−1
9	−1	−1	−1	−1	−1	19	−1	+1	−1	+1	+1
9	+1	+1	−1	+1	+1	19	+1	−1	+1	−1	+1
10	−1	−1	−1	+1	+1	20	−1	+1	+1	−1	+1
10	+1	+1	−1	−1	−1	20	+1	−1	−1	+1	+1

Table A7: Smallest optimal variance-balanced conjoint design with sets of $m = 4$ profiles, $k_c = 1$ constant attribute and $k_v = 4$ non-constant attributes.

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

Table A8: Smallest optimal variance-balanced conjoint design with sets of $m = 8$ profiles, $k_c = 1$ constant attribute and $k_v = 4$ non-constant attributes.

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

A.2. Designs with more than one constant attribute: $k_c > 1$

Table A9: Smallest optimal variance-balanced conjoint design with sets of $m = 2$ profiles, $k_c = 2$ constant attributes and $k_v = 4$ non-constant attributes.

PBIBD			Constant attribute design						Conjoint design					
Attrs			Attributes						Attributes					
cst		Set	1	2	3	4	5	6	1	2	3	4	5	6
1	4	1	−1	—	—	−1	—	—	−1	−1	−1	−1	−1	−1
		1							−1	+1	+1	−1	+1	+1
		2	−1	—	—	+1	—	—	−1	−1	−1	+1	+1	+1
		2							−1	+1	+1	+1	−1	−1
		3	+1	—	—	−1	—	—	+1	−1	+1	−1	−1	+1
		3							+1	+1	−1	−1	+1	−1
		4	+1	—	—	+1	—	—	+1	−1	+1	+1	+1	−1
		4							+1	+1	−1	+1	−1	+1
2	5	5	—	−1	—	—	−1	—	−1	−1	−1	−1	−1	+1
		5							+1	−1	+1	+1	−1	−1
		6	—	−1	—	—	+1	—	−1	−1	−1	+1	+1	−1
		6							+1	−1	+1	−1	+1	+1
		7	—	+1	—	—	−1	—	−1	+1	+1	−1	−1	−1
		7							+1	+1	−1	+1	−1	+1
		8	—	+1	—	—	+1	—	−1	+1	+1	+1	+1	+1
		8							+1	+1	−1	−1	+1	−1
3	6	9	—	—	−1	—	—	−1	−1	−1	−1	−1	−1	−1
		9							+1	+1	−1	+1	+1	−1
		10	—	—	−1	—	—	+1	−1	−1	−1	+1	+1	+1
		10							+1	+1	−1	−1	−1	+1
		11	—	—	+1	—	—	−1	−1	+1	+1	−1	+1	−1
		11							+1	−1	+1	+1	−1	−1
		12	—	—	+1	—	—	+1	−1	+1	+1	+1	−1	+1
		12							+1	−1	+1	−1	+1	+1

Table A10: Smallest optimal variance-balanced conjoint design with sets of $m = 4$ profiles, $k_c = 2$ constant attributes and $k_v = 4$ non-constant attributes.

PBIBD		Constant attribute design						Conjoint design					
Attrs		Attributes						Attributes					
cst	Set	1	2	3	4	5	6	1	2	3	4	5	6
1 4	1	−1	—	—	−1	—	—	−1	−1	−1	−1	−1	−1
	1							−1	+1	+1	−1	+1	+1
	1							−1	−1	−1	−1	+1	+1
	1							−1	+1	+1	−1	−1	−1
	2	−1	—	—	+1	—	—	−1	−1	+1	+1	−1	+1
	2							−1	+1	−1	+1	+1	−1
	2							−1	−1	+1	+1	+1	−1
	2							−1	+1	−1	+1	−1	+1
	3	+1	—	—	−1	—	—	+1	−1	−1	−1	−1	+1
	3							+1	+1	+1	−1	+1	−1
	3							+1	−1	−1	−1	+1	−1
	3							+1	+1	+1	−1	−1	+1
	4	+1	—	—	+1	—	—	+1	−1	+1	+1	−1	−1
	4							+1	+1	−1	+1	+1	+1
	4							+1	−1	+1	+1	+1	+1
	4							+1	+1	−1	+1	−1	−1
⋮	⋮	⋮						⋮					
3 6	9	—	—	−1	—	—	−1	−1	−1	−1	−1	−1	−1
	9							+1	+1	−1	+1	+1	−1
	9							−1	−1	−1	+1	+1	−1
	9							+1	+1	−1	−1	−1	−1
	10	—	—	−1	—	—	+1	−1	+1	−1	−1	+1	+1
	10							+1	−1	−1	+1	−1	+1
	10							−1	+1	−1	+1	−1	+1
	10							+1	−1	−1	−1	+1	+1
	11	—	—	+1	—	—	−1	−1	−1	+1	−1	+1	−1
	11							+1	+1	+1	+1	−1	−1
	11							−1	−1	+1	+1	−1	−1
	11							+1	+1	+1	−1	+1	−1
	12	—	—	+1	—	—	+1	−1	+1	+1	−1	−1	+1
	12							+1	−1	+1	+1	+1	+1
	12							−1	+1	+1	+1	+1	+1
	12							+1	−1	+1	−1	−1	+1

Table A11: Smallest optimal variance-balanced conjoint design with sets of $m = 8$ profiles, $k_c = 2$ constant attributes and $k_v = 4$ non-constant attributes.

PBIBD			Constant attribute design						Conjoint design					
Attrs			Attributes						Attributes					
cst		Set	1	2	3	4	5	6	1	2	3	4	5	6
		1							−1	−1	−1	−1	−1	−1
		1							−1	+1	+1	−1	+1	+1
		1							−1	−1	−1	−1	+1	+1
		1							−1	+1	+1	−1	−1	−1
		1	−1	—	—	−1	—	—	−1	−1	+1	−1	−1	+1
		1							−1	+1	−1	−1	+1	−1
		1							−1	−1	+1	−1	+1	−1
		1							−1	+1	−1	−1	−1	+1
		2							−1	−1	−1	+1	−1	+1
		2							−1	+1	+1	+1	+1	−1
		2							−1	−1	−1	+1	+1	−1
		2							−1	+1	+1	+1	−1	+1
		2	−1	—	—	+1	—	—	−1	−1	+1	+1	−1	−1
		2							−1	+1	−1	+1	+1	+1
		2							−1	−1	+1	+1	+1	+1
1	4	2							−1	+1	−1	+1	−1	−1
		3							+1	−1	−1	−1	−1	−1
		3							+1	+1	+1	−1	+1	+1
		3							+1	−1	−1	−1	+1	+1
		3							+1	+1	+1	−1	−1	−1
		3	+1	—	—	−1	—	—	+1	−1	+1	−1	−1	+1
		3							+1	+1	−1	−1	+1	−1
		3							+1	−1	+1	−1	+1	−1
		3							+1	+1	−1	−1	−1	+1
		4							+1	−1	−1	+1	−1	+1
		4							+1	+1	+1	+1	+1	−1
		4							+1	−1	−1	+1	+1	−1
		4							+1	+1	+1	+1	−1	+1
		4	+1	—	—	+1	—	—	+1	−1	+1	+1	−1	−1
		4							+1	+1	−1	+1	+1	+1
		4							+1	−1	+1	+1	+1	+1
		4							+1	+1	−1	+1	−1	−1
⋮		⋮				⋮						⋮		

continued on next page

continued from previous page

PBIBD			Constant attribute design						Conjoint design					
Attrs			Attributes						Attributes					
cst		Set	1	2	3	4	5	6	1	2	3	4	5	6
3	6	9	—	—	−1	—	—	−1	−1	−1	−1	−1	−1	−1
		9							+1	+1	−1	+1	+1	−1
		9							−1	−1	−1	+1	+1	−1
		9							+1	+1	−1	−1	−1	−1
		9							−1	+1	−1	−1	+1	−1
		9							+1	−1	−1	+1	−1	−1
		9							−1	+1	−1	+1	−1	−1
		9							+1	−1	−1	−1	+1	−1
		10	—	—	−1	—	—	+1	−1	−1	−1	−1	+1	+1
		10							+1	+1	−1	+1	−1	+1
		10							−1	−1	−1	+1	−1	+1
		10							+1	+1	−1	−1	+1	+1
		10							−1	+1	−1	−1	−1	+1
		10							+1	−1	−1	+1	+1	+1
		10							−1	+1	−1	+1	+1	+1
		10							+1	−1	−1	−1	−1	+1
		11	—	—	+1	—	—	−1	−1	−1	+1	−1	−1	−1
		11							+1	+1	+1	+1	+1	−1
		11							−1	−1	+1	+1	+1	−1
		11							+1	+1	+1	−1	−1	−1
		11							−1	+1	+1	−1	+1	−1
		11							+1	−1	+1	+1	−1	−1
		11							−1	+1	+1	+1	−1	−1
		11							+1	−1	+1	−1	+1	−1
		12	—	—	+1	—	—	+1	−1	−1	+1	−1	+1	+1
		12							+1	+1	+1	+1	−1	+1
		12							−1	−1	+1	+1	−1	+1
		12							+1	+1	+1	−1	+1	+1
		12							−1	+1	+1	−1	−1	+1
		12							+1	−1	+1	+1	+1	+1
		12							−1	+1	+1	+1	+1	+1
		12							+1	−1	+1	−1	−1	+1

Table A12: Smallest optimal variance-balanced conjoint design with sets of $m = 2$ profiles, $k_c = 2$ constant attributes and $k_v = 2$ non-constant attributes.

PBIBD			Cst attribute design				Conjoint design			
Attrs			Attributes				Attributes			
cst		Set	1	2	3	4	1	2	3	4
1	2	1	−1	−1	—	—	−1	−1	−1	−1
		1					−1	−1	+1	+1
		2	−1	+1	—	—	−1	+1	−1	+1
		2					−1	+1	+1	−1
1	4	3	+1	—	—	−1	+1	−1	−1	−1
		3					+1	+1	+1	−1
		4	+1	—	—	+1	+1	−1	+1	+1
		4					+1	+1	−1	+1
2	3	5	—	−1	−1	—	−1	−1	−1	−1
		5					+1	−1	−1	+1
		6	—	+1	−1	—	−1	+1	−1	+1
		6					+1	+1	−1	−1
3	4	7	—	—	+1	−1	−1	−1	+1	−1
		7					+1	+1	+1	−1
		8	—	—	+1	+1	−1	+1	+1	+1
		8					+1	−1	+1	+1

Table A13: Smallest optimal variance-balanced conjoint design with sets of $m = 4$ profiles, $k_c = 2$ constant attributes and $k_v = 2$ non-constant attributes.

PBIBD			Cst attribute design				Conjoint design			
Attrs			Attributes				Attributes			
cst		Set	1	2	3	4	1	2	3	4
1	2	1	−1	−1	—	—	−1	−1	−1	−1
		1					−1	−1	+1	+1
		1					−1	−1	−1	+1
		1					−1	−1	+1	−1
		2	−1	+1	—	—	−1	+1	−1	−1
		2					−1	+1	+1	+1
		2					−1	+1	−1	+1
		2					−1	+1	+1	−1
1	4	3	+1	—	—	−1	+1	−1	−1	−1
		3					+1	+1	+1	−1
		3					+1	−1	+1	−1
		3					+1	+1	−1	−1
		4	+1	—	—	+1	+1	−1	−1	+1
		4					+1	+1	+1	+1
		4					+1	−1	+1	+1
		4					+1	+1	−1	+1
2	3	5	—	−1	−1	—	−1	−1	−1	−1
		5					+1	−1	−1	+1
		5					−1	−1	−1	+1
		5					+1	−1	−1	−1
		6	—	+1	−1	—	−1	+1	−1	−1
		6					+1	+1	−1	+1
		6					−1	+1	−1	+1
		6					+1	+1	+1	−1
3	4	7	—	—	+1	−1	−1	−1	+1	−1
		7					+1	+1	+1	−1
		7					−1	+1	+1	−1
		7					+1	−1	+1	−1
		8	—	—	+1	+1	−1	−1	+1	+1
		8					+1	+1	+1	+1
		8					−1	+1	+1	+1
		8					+1	−1	+1	+1

Table A14: Smallest optimal variance-balanced conjoint design with sets of $m = 2$ profiles, $k_c = 3$ constant attributes and $k_v = 3$ non-constant attributes.

PBIBD				Constant attribute design						Conjoint design					
Attrs				Attributes						Attributes					
cst			Set	1	2	3	4	5	6	1	2	3	4	5	6
1	2	3	1	−1	−1	+1	—	—	—	−1	−1	+1	−1	−1	−1
			1							−1	−1	+1	+1	+1	+1
			2	−1	+1	−1	—	—	—	−1	+1	−1	−1	−1	+1
			2							−1	+1	−1	+1	+1	−1
			3	+1	−1	−1	—	—	—	+1	−1	−1	−1	+1	−1
			3							+1	−1	−1	+1	−1	+1
			4	+1	+1	+1	—	—	—	+1	+1	+1	−1	+1	+1
			4							+1	+1	+1	+1	−1	−1
1	5	6	5	−1	—	—	—	−1	−1	−1	−1	−1	−1	−1	−1
			5							−1	+1	+1	+1	−1	−1
			6	−1	—	—	—	+1	+1	−1	−1	−1	+1	+1	+1
			6							−1	+1	+1	−1	+1	+1
			7	+1	—	—	—	−1	+1	+1	−1	+1	−1	−1	+1
			7							+1	+1	−1	+1	−1	+1
			8	+1	—	—	—	+1	−1	+1	−1	+1	+1	+1	−1
			8							+1	+1	−1	−1	+1	−1
2	4	6	9	—	−1	—	−1	—	+1	−1	−1	−1	−1	−1	+1
			9							+1	−1	+1	−1	+1	+1
			10	—	−1	—	+1	—	−1	−1	−1	−1	+1	+1	−1
			10							+1	−1	+1	+1	−1	−1
			11	—	+1	—	−1	—	−1	−1	+1	+1	−1	−1	−1
			11							+1	+1	−1	−1	+1	−1
			12	—	+1	—	+1	—	+1	−1	+1	+1	+1	+1	+1
			12							+1	+1	−1	+1	−1	+1
3	4	5	13	—	—	−1	−1	−1	—	−1	−1	−1	−1	−1	−1
			13							+1	+1	−1	−1	−1	+1
			14	—	—	−1	+1	+1	—	−1	−1	−1	+1	+1	+1
			14							+1	+1	−1	+1	+1	−1
			15	—	—	+1	−1	+1	—	−1	+1	+1	−1	+1	−1
			15							+1	−1	+1	−1	+1	+1
			16	—	—	+1	+1	−1	—	−1	+1	+1	+1	−1	+1
			16							+1	−1	+1	+1	−1	−1

Table A15: Smallest optimal variance-balanced conjoint design with sets of $m = 2$ profiles, $k_c = 4$ constant attributes and $k_v = 4$ non-constant attributes.

PBIBD		Constant attribute design								Conjoint design							
Attrs		Attributes								Attributes							
cst	Set	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
1 2 5 6	1	−1	−1	—	—	−1	−1	—	—	−1	−1	−1	−1	−1	−1	−1	−1
	1									−1	−1	+1	+1	−1	−1	+1	+1
	2	−1	−1	—	—	+1	+1	—	—	−1	−1	−1	−1	+1	+1	+1	+1
	2									−1	−1	+1	+1	+1	+1	−1	−1
	3	−1	+1	—	—	−1	+1	—	—	−1	+1	−1	+1	−1	+1	−1	+1
	3									−1	+1	+1	−1	−1	+1	+1	−1
	4	−1	+1	—	—	+1	−1	—	—	−1	+1	−1	+1	+1	−1	+1	−1
	4									−1	+1	+1	−1	+1	−1	−1	+1
	5	+1	−1	—	—	−1	+1	—	—	+1	−1	−1	−1	−1	+1	−1	+1
	5									+1	−1	+1	+1	−1	+1	+1	−1
	6	+1	−1	—	—	+1	−1	—	—	+1	−1	−1	−1	+1	−1	+1	−1
	6									+1	−1	+1	+1	+1	−1	−1	+1
	7	+1	+1	—	—	−1	−1	—	—	+1	+1	−1	+1	−1	−1	−1	−1
	7									+1	+1	+1	−1	−1	−1	+1	+1
	8	+1	+1	—	—	+1	+1	—	—	+1	+1	−1	+1	+1	+1	+1	+1
	8									+1	+1	+1	−1	+1	+1	−1	−1
1 3 5 7	9	−1	—	−1	—	−1	—	+1	—	−1	−1	−1	−1	−1	−1	+1	−1
	9									−1	+1	−1	+1	−1	+1	+1	+1
	10	−1	—	−1	—	+1	—	−1	—	−1	−1	−1	−1	+1	+1	−1	+1
	10									−1	+1	−1	+1	+1	−1	−1	−1
	11	−1	—	+1	—	−1	—	−1	—	−1	−1	+1	+1	−1	−1	−1	+1
	11									−1	+1	+1	−1	−1	+1	−1	−1
	12	−1	—	+1	—	+1	—	+1	—	−1	−1	+1	+1	+1	+1	+1	−1
	12									−1	+1	+1	−1	+1	−1	+1	+1
	13	+1	—	−1	—	−1	—	−1	—	+1	−1	−1	−1	−1	−1	−1	+1
	13									+1	+1	−1	+1	−1	+1	−1	−1
	14	+1	—	−1	—	+1	—	+1	—	+1	−1	−1	−1	+1	+1	+1	−1
	14									+1	+1	−1	+1	+1	−1	+1	+1
	15	+1	—	+1	—	−1	—	+1	—	+1	−1	+1	+1	−1	−1	+1	−1
	15									+1	+1	+1	−1	−1	+1	+1	+1
	16	+1	—	+1	—	+1	—	−1	—	+1	−1	+1	+1	+1	+1	−1	+1
	16									+1	+1	+1	−1	+1	−1	−1	−1

continued on next page

continued from previous page

PBIBD				Constant attribute design								Conjoint design							
Attrs					Attributes							Attributes							
cst				Set	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
1	4	5	8	17	−1	—	—	−1	−1	—	—	−1	−1	−1	−1	−1	−1	−1	−1	−1
				17									−1	+1	+1	−1	−1	+1	+1	−1
				18	−1	—	—	−1	+1	—	—	+1	−1	−1	−1	−1	+1	+1	+1	+1
				18									−1	+1	+1	−1	+1	−1	−1	+1
				19	−1	—	—	+1	−1	—	—	+1	−1	−1	+1	+1	−1	−1	+1	+1
				19									−1	+1	−1	+1	−1	+1	−1	+1
				20	−1	—	—	+1	+1	—	—	−1	−1	−1	+1	+1	+1	+1	−1	−1
				20									−1	+1	−1	+1	+1	−1	+1	−1
				21	+1	—	—	−1	−1	—	—	+1	+1	−1	−1	−1	−1	−1	+1	+1
				21									+1	+1	+1	−1	−1	+1	−1	+1
				22	+1	—	—	−1	+1	—	—	−1	+1	−1	−1	−1	+1	+1	−1	−1
				22									+1	+1	+1	−1	+1	−1	+1	−1
				23	+1	—	—	+1	−1	—	—	−1	+1	−1	+1	+1	−1	−1	−1	−1
				23									+1	+1	−1	+1	−1	+1	+1	−1
				24	+1	—	—	+1	+1	—	—	+1	+1	−1	+1	+1	+1	+1	+1	+1
				24									+1	+1	−1	+1	+1	−1	−1	+1
2	3	6	7	25	—	−1	−1	—	—	−1	+1	—	−1	−1	−1	−1	−1	−1	+1	−1
				25									+1	−1	−1	+1	+1	−1	+1	+1
				26	—	−1	−1	—	—	+1	−1	—	−1	−1	−1	−1	+1	+1	−1	+1
				26									+1	−1	−1	+1	−1	+1	−1	−1
				27	—	−1	+1	—	—	−1	−1	—	−1	−1	+1	+1	−1	−1	−1	+1
				27									+1	−1	+1	−1	+1	−1	−1	−1
				28	—	−1	+1	—	—	+1	+1	—	−1	−1	+1	+1	+1	+1	+1	−1
				28									+1	−1	+1	−1	−1	+1	+1	+1
				29	—	+1	−1	—	—	−1	−1	—	−1	+1	−1	−1	−1	−1	−1	+1
				29									+1	+1	−1	+1	+1	−1	−1	−1
				30	—	+1	−1	—	—	+1	+1	—	−1	+1	−1	−1	+1	+1	+1	−1
				30									+1	+1	−1	+1	−1	+1	+1	+1
				31	—	+1	+1	—	—	−1	+1	—	−1	+1	+1	+1	−1	−1	+1	−1
				31									+1	+1	+1	−1	+1	−1	+1	+1
				32	—	+1	+1	—	—	+1	−1	—	−1	+1	+1	+1	+1	+1	−1	+1
				32									+1	+1	+1	−1	−1	+1	−1	−1

continued on next page

continued from previous page

PBIBD		Constant attribute design								Conjoint design							
Attrs		Attributes								Attributes							
cst	Set	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
	33	—	−1	—	−1	—	−1	—	−1	−1	−1	−1	−1	−1	−1	−1	−1
	33	—	−1	—	−1	—	−1	—	−1	+1	−1	+1	−1	+1	−1	+1	−1
	34	—	−1	—	−1	—	+1	—	+1	−1	−1	−1	−1	+1	+1	+1	+1
	34	—	−1	—	−1	—	+1	—	+1	+1	−1	+1	−1	−1	+1	−1	+1
	35	—	−1	—	+1	—	−1	—	+1	−1	−1	+1	+1	−1	−1	+1	+1
	35	—	−1	—	+1	—	−1	—	+1	+1	−1	−1	+1	+1	−1	−1	+1
	36	—	−1	—	+1	—	+1	—	−1	−1	−1	+1	+1	+1	+1	−1	−1
	36	—	−1	—	+1	—	+1	—	−1	+1	−1	−1	+1	−1	+1	+1	−1
2	37	—	+1	—	−1	—	−1	—	+1	−1	+1	−1	−1	−1	−1	+1	+1
4	37	—	+1	—	−1	—	−1	—	+1	+1	+1	+1	−1	+1	−1	−1	+1
6	38	—	+1	—	−1	—	+1	—	−1	−1	+1	−1	+1	+1	+1	−1	−1
8	38	—	+1	—	−1	—	+1	—	−1	+1	+1	+1	−1	−1	+1	+1	−1
	39	—	+1	—	+1	—	−1	—	−1	−1	+1	+1	+1	−1	−1	−1	−1
	39	—	+1	—	+1	—	−1	—	−1	+1	+1	−1	+1	−1	−1	+1	−1
	40	—	+1	—	+1	—	+1	—	+1	−1	+1	+1	+1	+1	+1	+1	+1
	40	—	+1	—	+1	—	+1	—	+1	+1	+1	+1	+1	−1	+1	−1	+1
	41	—	—	−1	−1	—	—	−1	+1	−1	−1	−1	−1	−1	−1	−1	+1
	41	—	—	−1	−1	—	—	−1	+1	+1	+1	−1	−1	+1	+1	+1	−1
	42	—	—	−1	−1	—	—	+1	−1	−1	−1	−1	+1	−1	+1	+1	−1
	42	—	—	−1	−1	—	—	+1	−1	+1	+1	−1	−1	−1	−1	+1	−1
	43	—	—	−1	+1	—	—	−1	−1	−1	+1	−1	+1	−1	+1	−1	−1
	43	—	—	−1	+1	—	—	−1	−1	+1	−1	−1	+1	+1	−1	−1	−1
	44	—	—	−1	+1	—	—	+1	+1	−1	+1	−1	+1	−1	+1	+1	+1
	44	—	—	−1	+1	—	—	+1	+1	+1	−1	−1	+1	+1	−1	+1	+1
	45	—	—	+1	−1	—	—	−1	−1	−1	−1	+1	−1	−1	+1	−1	−1
	45	—	—	+1	−1	—	—	−1	−1	+1	+1	+1	−1	+1	−1	−1	−1
	46	—	—	+1	−1	—	—	+1	+1	−1	−1	+1	−1	+1	−1	+1	+1
	46	—	—	+1	−1	—	—	+1	+1	+1	+1	+1	−1	−1	+1	+1	+1
	47	—	—	+1	+1	—	—	−1	+1	−1	+1	+1	+1	−1	−1	−1	+1
	47	—	—	+1	+1	—	—	−1	+1	+1	−1	+1	+1	+1	+1	−1	+1
	48	—	—	+1	+1	—	—	+1	−1	−1	+1	+1	+1	+1	+1	+1	−1
	48	—	—	+1	+1	—	—	+1	−1	+1	−1	+1	+1	−1	−1	−1	−1

Appendix B. Derivation of the information matrix

In this appendix, we derive the information matrix of the optimal variance-balanced conjoint design in Table A12 of Appendix A.2. Based on this example, we provide a general method to compute the information matrix of any optimal variance-balanced conjoint design.

The conjoint design in Table A12 has $b = 8$ sets of $m = 2$ profiles, $k_c = 2$ constant attributes and $k_v = 2$ non-constant attributes. It is built on a PBIBD that identifies four combinations of constant attributes. Each attribute is constant in four profile sets. To compute the information matrix of the design, we go back to the work of Goos and Vandebroek (2001b) which provides the basis for the information matrix (5.8). Under assumptions (5.3), (5.4) and (5.5), $\mathbf{V}$ is block diagonal, so that

$$\mathbf{X}'\mathbf{V}^{-1}\mathbf{X} = \sum_{i=1}^{b} \mathbf{X}_i'\mathbf{V}_m^{-1}\mathbf{X}_i, \tag{B1}$$

where

$$\mathbf{V}_m^{-1} = \sigma_{\boldsymbol{\varepsilon}}^{-2}\left(\mathbf{I}_m - \frac{\rho}{1+\rho(m-1)}\mathbf{1}_m\mathbf{1}_m'\right). \tag{B2}$$

According to (B1), the information matrix of the design in Table A12 is the sum of $b = 8$ information matrices, one for each of the profile sets. We now compute the information matrix for each profile set and sum the matrices to obtain the total information matrix of the design. The design has profile sets of size $m = 2$ so that

$$\mathbf{V}_2^{-1} = \sigma_{\boldsymbol{\varepsilon}}^{-2}\left(\mathbf{I}_2 - \frac{\rho}{1+\rho}\mathbf{1}_2\mathbf{1}_2'\right). \tag{B3}$$

To derive the information matrices of the individual profile sets, we write each profile set in terms of its constant and non-constant attributes. Therefore, we denote the levels of the two constant attributes in profile set i by $w_{i1}\mathbf{1}_2$ and $w_{i2}\mathbf{1}_2$, where w_{i1} and w_{i2} equal either -1 and $+1$, and we denote the levels of the two non-constant attributes by the two-dimensional vectors $\mathbf{s}_{i1}$ and $\mathbf{s}_{i2}$. Because Profile sets 1 and 2 have the first two attributes constant,

we can write them as $\mathbf{X}_i = [\mathbf{1}_2, w_{i1}\mathbf{1}_2, w_{i2}\mathbf{1}_2, \mathbf{s}_{i1}, \mathbf{s}_{i2}]$, where $i = 1, 2$. The information matrix for each of the two profile sets then is

$$\begin{aligned}
&\mathbf{X}_i'\mathbf{V}_2^{-1}\mathbf{X}_i \\
&\quad= \begin{bmatrix}
\mathbf{1}_2'\mathbf{V}_2^{-1}\mathbf{1}_2 & w_{i1}\mathbf{1}_2'\mathbf{V}_2^{-1}\mathbf{1}_2 & w_{i2}\mathbf{1}_2'\mathbf{V}_2^{-1}\mathbf{1}_2 & \mathbf{1}_2'\mathbf{V}_2^{-1}\mathbf{s}_{i1} & \mathbf{1}_2'\mathbf{V}_2^{-1}\mathbf{s}_{i2} \\
w_{i1}\mathbf{1}_2'\mathbf{V}_2^{-1}\mathbf{1}_2 & w_{i1}^2\mathbf{1}_2'\mathbf{V}_2^{-1}\mathbf{1}_2 & w_{i1}w_{i2}\mathbf{1}_2'\mathbf{V}_2^{-1}\mathbf{1}_2 & w_{i1}\mathbf{1}_2'\mathbf{V}_2^{-1}\mathbf{s}_{i1} & w_{i1}\mathbf{1}_2'\mathbf{V}_2^{-1}\mathbf{s}_{i2} \\
w_{i2}\mathbf{1}_2'\mathbf{V}_2^{-1}\mathbf{1}_2 & w_{i1}w_{i2}\mathbf{1}_2'\mathbf{V}_2^{-1}\mathbf{1}_2 & w_{i2}^2\mathbf{1}_2'\mathbf{V}_2^{-1}\mathbf{1}_2 & w_{i2}\mathbf{1}_2'\mathbf{V}_2^{-1}\mathbf{s}_{i1} & w_{i2}\mathbf{1}_2'\mathbf{V}_2^{-1}\mathbf{s}_{i2} \\
\mathbf{s}_{i1}'\mathbf{V}_2^{-1}\mathbf{1}_2 & w_{i1}\mathbf{s}_{i1}'\mathbf{V}_2^{-1}\mathbf{1}_2 & w_{i2}\mathbf{s}_{i1}'\mathbf{V}_2^{-1}\mathbf{1}_2 & \mathbf{s}_{i1}'\mathbf{V}_2^{-1}\mathbf{s}_{i1} & \mathbf{s}_{i1}'\mathbf{V}_2^{-1}\mathbf{s}_{i2} \\
\mathbf{s}_{i2}'\mathbf{V}_2^{-1}\mathbf{1}_2 & w_{i1}\mathbf{s}_{i2}'\mathbf{V}_2^{-1}\mathbf{1}_2 & w_{i2}\mathbf{s}_{i2}'\mathbf{V}_2^{-1}\mathbf{1}_2 & \mathbf{s}_{i2}'\mathbf{V}_2^{-1}\mathbf{s}_{i1} & \mathbf{s}_{i2}'\mathbf{V}_2^{-1}\mathbf{s}_{i2}
\end{bmatrix}, \\
&\quad= \sigma_\varepsilon^{-2}\begin{bmatrix}
2c_2 & 2w_{i1}c_2 & 2w_{i2}c_2 & c_2\mathbf{1}_2'\mathbf{s}_{i1} & c_2\mathbf{1}_2'\mathbf{s}_{i2} \\
2w_{i1}c_2 & 2c_2 & 2w_{i1}w_{i2}c_2 & w_{i1}c_2\mathbf{1}_2'\mathbf{s}_{i1} & w_{i1}c_2\mathbf{1}_2'\mathbf{s}_{i2} \\
2w_{i2}c_2 & 2w_{i1}w_{i2}c_2 & 2c_2 & w_{i2}c_2\mathbf{1}_2'\mathbf{s}_{i1} & w_{i2}c_2\mathbf{1}_2'\mathbf{s}_{i2} \\
c_2\mathbf{s}_{i1}'\mathbf{1}_2 & w_{i1}c_2\mathbf{s}_{i1}'\mathbf{1}_2 & w_{i2}c_2\mathbf{s}_{i1}'\mathbf{1}_2 & \mathbf{s}_{i1}'\mathbf{s}_{i1} - t_{11} & \mathbf{s}_{i1}'\mathbf{s}_{i2} - t_{12} \\
c_2\mathbf{s}_{i2}'\mathbf{1}_2 & w_{i1}c_2\mathbf{s}_{i2}'\mathbf{1}_2 & w_{i2}c_2\mathbf{s}_{i2}'\mathbf{1}_2 & \mathbf{s}_{i2}'\mathbf{s}_{i1} - t_{21} & \mathbf{s}_{i2}'\mathbf{s}_{i2} - t_{22}
\end{bmatrix},
\end{aligned} \tag{B4}$$

where $c_2 = (1-\rho)/(1+\rho)$, $t_{11} = \frac{\rho}{1-\rho}c_2(\mathbf{s}_{i1}'\mathbf{1}_2)(\mathbf{1}_2'\mathbf{s}_{i1})$, $t_{12} = \frac{\rho}{1-\rho}c_2(\mathbf{s}_{i1}'\mathbf{1}_2)(\mathbf{1}_2'\mathbf{s}_{i2})$, and t_{21} and t_{22} are obtained similarly to t_{12} and t_{11}, respectively. Because the levels of the non-constant attributes are balanced in the profile sets, $\mathbf{s}_{i1}'\mathbf{1}_2 = \mathbf{1}_2'\mathbf{s}_{i1} = \mathbf{s}_{i2}'\mathbf{1}_2 = \mathbf{1}_2'\mathbf{s}_{i2} = 0$. As a result, filling out the entries of the formal information matrix (B4) for profile set $i = 1$ yields the information matrix

$$\mathbf{X}_1'\mathbf{V}_2^{-1}\mathbf{X}_1 = \sigma_\varepsilon^{-2}\begin{bmatrix}
+2c_2 & -2c_2 & -2c_2 & 0 & 0 \\
-2c_2 & +2c_2 & +2c_2 & 0 & 0 \\
-2c_2 & +2c_2 & +2c_2 & 0 & 0 \\
0 & 0 & 0 & +2 & +2 \\
0 & 0 & 0 & +2 & +2
\end{bmatrix}, \tag{B5}$$

and for profile set $i = 2$ the information matrix is

$$\mathbf{X}_2'\mathbf{V}_2^{-1}\mathbf{X}_2 = \sigma_\varepsilon^{-2}\begin{bmatrix}
+2c_2 & -2c_2 & +2c_2 & 0 & 0 \\
-2c_2 & +2c_2 & -2c_2 & 0 & 0 \\
+2c_2 & -2c_2 & +2c_2 & 0 & 0 \\
0 & 0 & 0 & +2 & -2 \\
0 & 0 & 0 & -2 & +2
\end{bmatrix}. \tag{B6}$$

Summing these two matrices results in

$$\sum_{i=1}^{2}\mathbf{X}_i'\mathbf{V}_2^{-1}\mathbf{X}_i = \sigma_\varepsilon^{-2}\begin{bmatrix}
+4c_2 & -4c_2 & 0 & 0 & 0 \\
-4c_2 & +4c_2 & 0 & 0 & 0 \\
0 & 0 & +4c_2 & 0 & 0 \\
0 & 0 & 0 & +4 & 0 \\
0 & 0 & 0 & 0 & +4
\end{bmatrix}. \tag{B7}$$

The diagonal elements for the part-worths in (B7) reveal that, since $c_2 < 1$, less information is obtained on the constant attributes than on the non-constant attributes. If we had not included the constant attributes in the profile sets, the diagonal elements for the part-worths would have been all equal to four, the number of profiles in the two profile sets. Also, the two non-diagonal elements would have vanished.

The next pairs of profile sets in the design of Table A12 have Attributes 1 and 4, 2 and 3, and 3 and 4 as constant attributes. Hence, we denote Profile sets 3 and 4 as $\mathbf{X}_i = [\mathbf{1}_2, w_{i1}\mathbf{1}_2, \mathbf{s}_{i1}, \mathbf{s}_{i2}, w_{i2}\mathbf{1}_2]$, where $i = 3, 4$. We denote Profile sets 5 and 6 as $\mathbf{X}_i = [\mathbf{1}_2, \mathbf{s}_{i1}, w_{i1}\mathbf{1}_2, w_{i2}\mathbf{1}_2, \mathbf{s}_{i2}]$, where $i = 5, 6$ and we denote Profile sets 7 and 8 as $\mathbf{X}_i = [\mathbf{1}_2, \mathbf{s}_{i1}, \mathbf{s}_{i2}, w_{i1}\mathbf{1}_2, w_{i2}\mathbf{1}_2]$, where $i = 7, 8$. By the same procedure as described above, the information matrices for each of these pairs are

$$\sum_{i=3}^{4} \mathbf{X}_i' \mathbf{V}_2^{-1} \mathbf{X}_i = \sigma_{\varepsilon}^{-2} \begin{bmatrix} +4c_2 & +4c_2 & 0 & 0 & 0 \\ +4c_2 & +4c_2 & 0 & 0 & 0 \\ 0 & 0 & +4 & 0 & 0 \\ 0 & 0 & 0 & +4 & 0 \\ 0 & 0 & 0 & 0 & +4c_2 \end{bmatrix}, \tag{B8}$$

$$\sum_{i=5}^{6} \mathbf{X}_i' \mathbf{V}_2^{-1} \mathbf{X}_i = \sigma_{\varepsilon}^{-2} \begin{bmatrix} +4c_2 & 0 & 0 & -4c_2 & 0 \\ 0 & +4 & 0 & 0 & 0 \\ 0 & 0 & +4c_2 & 0 & 0 \\ -4c_2 & 0 & 0 & +4c_2 & 0 \\ 0 & 0 & 0 & 0 & +4 \end{bmatrix} \text{ and} \tag{B9}$$

$$\sum_{i=7}^{8} \mathbf{X}_i' \mathbf{V}_2^{-1} \mathbf{X}_i = \sigma_{\varepsilon}^{-2} \begin{bmatrix} +4c_2 & 0 & 0 & +4c_2 & 0 \\ 0 & +4 & 0 & 0 & 0 \\ 0 & 0 & +4 & 0 & 0 \\ +4c_2 & 0 & 0 & +4c_2 & 0 \\ 0 & 0 & 0 & 0 & +4c_2 \end{bmatrix}. \tag{B10}$$

Summing over the information matrices (B7), (B8), (B9) and (B10) for the four pairs of profile sets yields the total information matrix of the conjoint

design:

$$\begin{aligned}\mathbf{X}'\mathbf{V}^{-1}\mathbf{X} &= \sum_{i=1}^{8}\mathbf{X}_i'\mathbf{V}_2^{-1}\mathbf{X}_i, \\ &= \sigma_{\boldsymbol{\epsilon}}^{-2}\text{diag}[16c_2 \;\; 8c_2+8 \;\; 8c_2+8 \;\; 8c_2+8 \;\; 8c_2+8].\end{aligned} \tag{B11}$$

The amount of information on each of the part-worths is thus the same so that the conjoint design in Table A12 is variance-balanced. The information component $8c_2$ for each part-worth refers to the information on each attribute when it is constant. In the design each attribute is constant in 8 profiles. The other information component, 8, for each part-worth points at the information on each attribute when it is non-constant. Each attribute is also non-constant in 8 profiles. If there were no constant attributes in the design, the diagonal elements for the part-worths would have been equal to 16, the number of profiles in the design. The information on the intercept always amounts to $16c_2$.

An information matrix similar to the one in (B11) can be derived for any of the optimal variance-balanced conjoint designs presented in this chapter. In the cases $m = 4$ and $m = 8$, the value for c_2 in the formulas is no longer appropriate and needs to be replaced by $c_4 = (1-\rho)/(1+3\rho)$ and $c_8 = (1-\rho)/(1+7\rho)$, respectively. Note that $c_8 < c_4 < c_2 < 1$. In general, to obtain the amount of information on each part-worth, two information components must be computed because of the partitioning in constant and non-constant attributes. The first information component contains c_m multiplied by the number of profiles in which each attribute is constant. The second information component equals the number of profiles in which each attribute is non-constant. The amount of information on the intercept is obtained by multiplying the sample size by c_m.

List of Figures

List of Tables

Bibliography

Abel, R. J. R. and Greig, M. (1996). BIBDs with small block size, in *The CRC Handbook of Combinatorial Designs*, Ch. I.2, Colbourn, C. J. and Dinitz, J. H., eds. Boca Raton (Florida): CRC Press, 41–47.

Anderson, D. A. and Wiley, J. B. (1992). Efficient choice set designs for estimating availability cross-effects models, *Marketing Letters* **3**: 357–370.

Andrews, R. L., Ainslie, A. and Currim, I. S. (2002). An empirical comparison of logit choice models with discrete versus continuous representations of heterogeneity, *Journal of Marketing Research* **39**: 479–487.

Arora, N. and Huber, J. (2001). Improving parameter estimates and model prediction by aggregate customization in choice experiments, *Journal of Consumer Research* **28**: 273–283.

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

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

Bennett, J. M. (1965). Triangular factors of modified matrices, *Numerische Mathematik* **7**: 217–221.

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.

Bunch, D. S., Louviere, J. J. and Anderson, D. A. (1996). A comparison of experimental design strategies for multinomial logit models: the case of generic attributes, working paper, Graduate School of Management, University of California, Davis.

Burgess, L. and Street, D. J. (2003). Optimal designs for 2^k choice experiments, *Communications in Statistics: Theory and Methods* **32**: 2185–2206.

Burgess, L. and Street, D. J. (2005). Optimal designs for choice experiments with asymmetric attributes, *Journal of Statistical Planning and Inference*, **134**: 288–301.

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

Carson, R. T., Louviere, J. J., Anderson, D. A., Arabie, P., Bunch, D. S., Hensher, D. A., Johnson, R. M., Kuhfeld, W.F., Steinberg, D., Swait, J., Timmermans, H. and Wiley, J. B. (1994). Experimental analysis of choice, *Marketing Letters* **5**: 351–367.

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

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

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

Cox, D. R. (1958). *Planning of Experiments*, New York: Wiley.

Cook, R. D. and Nachtsheim, C. J. (1980). A comparison of algorithms for constructing exact $\mathcal{D}$-optimal designs, *Technometrics* **22**: 315–324.

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

Entholzner, M., Benda, N., Schmelter, T. and Schwabe, R. (2005). A note on designs for estimating population parameters, *Biometrical Letters – Listy Biometryczne* **42**: 25–41.

Fedorov, V. V. (1972). *Theory of Optimal Experiments*, New York: Academic Press.

Geweke, J. (1996). Monte Carlo simulation and numerical integration, in *Handbook of Computational Economics*, Vol. 1, Ch. 15, Amman H. M., Kendrick D. A., Rust J., eds. Amsterdam: Elsevier Science, 731–800.

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. (2006a). Optimal versus orthogonal and equivalent-estimation design of blocked and split-plot experiments, *Statistica Neerlandica* **60**: 361–378.

Goos, P. (2006b). Optimal split-plot orthogonal main-effects plans, working paper, Department of Mathematics, Statistics and Actuarial Sciences, Faculty of Applied Economics, Universiteit Antwerpen, to appear.

Goos, P. and Vandebroek, M. (2001a). $\mathcal{D}$-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.

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

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

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

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

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

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

Hickernell, F. J., Hong, H. S., L'Ecuyer, P. and Lemieux, C. (2000). Extensible lattice sequences for quasi-Monte Carlo quadrature, *SIAM Journal on Scientific Computing* **22**: 1117–1138.

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

Kessels, R., Goos, P. and Vandebroek, M. (2004). Optimal designs for rating-based conjoint experiments, working paper, Katholieke Universiteit Leuven, Belgium.

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

Kessels, R., Jones, B., Goos, P. and Vandebroek, M. (2006). An efficient algorithm for constructing Bayesian optimal choice designs, working paper, Katholieke Universiteit Leuven, Belgium.

Kuhfeld, W. F. and Tobias, R. D. (2005). Large factorial designs for product engineering and marketing research applications, *Technometrics* **47**: 132–141.

Kuhfeld, W. F., Tobias, R. D. and Garratt, M. (1994). Efficient experimental design with marketing research applications, *Journal of Marketing Research* **31**: 545–557. Updated version [2005] available at http://support.sas.com/techsup/tnote/tnote_stat.html.

Lazari, A. G. and Anderson, D. A. (1994). Designs of discrete choice experiments for estimating both attribute and availability cross effects, *Journal of Marketing Research* **31**: 375–383.

Lenk, P. J., DeSarbo, W. S., Green, P. E. and Young, M. R. (1996). Hierarchical Bayes conjoint analysis: recovery of part-worth 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.

Louviere, J. J., Street, D. J. and Burgess, L. (2003). A 20+ years retrospective on choice experiments, in *Marketing Research and Modeling: Progress and Prospects*, Ch. 8, Wind, Y. and Green, P. E., eds. New York: Kluwer Academic Press, 201–214.

Louviere, J. J. and Woodworth, G. (1983). Design and analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data, *Journal of Marketing Research* **20**: 350–367.

Mathon, R. and Rosa, A. (1996). 2-(v, k, λ) designs of small order, in *The CRC Handbook of Combinatorial Designs*, Ch. I.1, Colbourn, C. J. and Dinitz, J. H., eds. Boca Raton (Florida): CRC Press, 3–41.

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

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

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

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

Sándor, Z. and András, P. (2004). Alternative sampling methods for estimating multivariate normal probabilities, *Journal of Econometrics* **120**: 207–234.

Sándor, Z. and Train, K. (2004). Quasi-random simulation of discrete choice models, *Transportation Research B* **38**: 313–327.

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

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

Sándor, Z. and Wedel, M. (2005). Heterogeneous conjoint choice designs, *Journal of Marketing Research* **42**: 210–218.

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

Severin, V. (2000). Comparing statistical efficiency and respondent efficiency in choice experiments, PhD Thesis, University of Sidney.

Shah, K. R. and Sinha, B. K. (1989). *Theory of Optimal Designs*, New York: Springer-Verlag.

Silvey, S. D. (1980). *Optimal Design*, London: Chapman and Hall.

Sinha, K. Website of partially balanced incomplete block designs: http://www.geocities.com/kishsinha2000/partiallybalanceddesigns.html.

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* **30**: 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* **118**: 185–199.

Street, D. J., Burgess, L. and Louviere, J. J. (2005). Quick and easy choice sets: constructing optimal and nearly optimal stated choice experiments, *International Journal of Research in Marketing* **22**: 459–470.

Street, D. J. and Street, A. P. (1996). Partially balanced incomplete block designs, in *The CRC Handbook of Combinatorial Designs*, Colbourn, C. J. and Dinitz, J. H., Ch. IV.35, eds. Boca Raton (Florida): CRC Press, 419–423.

Swait, J. and Adamowicz, W. (2001). The influence of task complexity on consumer choice: a latent class model of decision strategy switching, *Journal of Consumer Research* **28**: 135–148.

Tang, B. (1993). Orthogonal array-based Latin hypercubes, *Journal of the American Statistical Association* **88**: 1392–1397.

Train, K. E. (2003). *Discrete Choice Methods with Simulation*, Cambridge U.K.: Cambridge University Press.

Trinca, L. A. and Gilmour, S. G. (2000). An algorithm for arranging response surface designs in small blocks, *Computational Statistics and Data Analysis* **33**: 25–43.

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.

Zwerina, K., Huber, J. and Kuhfeld, W. F. (1996). A general method for constructing efficient choice designs, working paper, Fuqua School of Business, Duke University, Durham, NC 27708. Updated version [2005] available at http://support.sas.com/techsup/tnote/tnote_stat.html.

Doctoral Dissertations from the Faculty of Economic and Applied Economic Sciences

From August 1, 1971.

1. GEPTS Stefaan (1971)
   Stability and efficiency of resource allocation processes in discrete commodity spaces. Leuven, K. U. Leuven, 1971. 86 pp.

2. PEETERS Theo (1971)
   Determinanten van de internationale handel in fabrikaten. Leuven, Acco, 1971. 290 pp.

3. VAN LOOY Wim (1971)
   Personeelsopleiding: een onderzoek naar investeringsaspekten van opleiding. Hasselt, Vereniging voor wetenschappelijk onderzoek in Limburg, 1971. VII, 238 pp.

4. THARAKAN Mathew (1972)
   Indian exports to the European community: problems and prospects. Leuven, Faculty of economics and applied economics, 1972. X, 343 pp.

5. HERROELEN Willy (1972)
   Heuristische programmatie: methodologische benadering en praktische toepassing op complexe combinatorische problemen. Leuven, Aurelia scientifica, 1972. X, 367 pp.

6. VANDENBULCKE Jacques (1973)
De studie en de evaluatie van data-organisatiemethodes en data-zoekmethodes. Leuven, s.n., 1973. 3 V.

7. PENNYCUICK Roy A. (1973)
The economics of the ecological syndrome. Leuven, Acco, 1973. XII, 177 pp.

8. KAWATA T. Bualum (1973)
Formation du capital d'origine belge, dette publique et stratégie du développement au Zaire. Leuven, K. U. Leuven, 1973. V, 342 pp.

9. DONCKELS Rik (1974)
Doelmatige oriëntering van de sectorale subsidiepolitiek in België: een theoretisch onderzoek met empirische toetsing. Leuven, K. U. Leuven, 1974. VII, 156 pp.

10. VERHELST Maurice (1974)
Contribution to the analysis of organizational information systems and their financial benefits. Leuven, K. U. Leuven, 1974. 2 V.

11. CLEMEUR Hugo (1974)
Enkele verzekeringstechnische vraagstukken in het licht van de nutstheorie. Leuven, Aurelia scientifica, 1974. 193 pp.

12. HEYVAERT Edward (1975)
De ontwikkeling van de moderne bank- en krediettechniek tijdens de zestiende en zeventiende eeuw in Europa en te Amsterdam in het bijzonder. Leuven, K. U. Leuven, 1975. 186 pp.

13. VERTONGHEN Robert (1975)
Investeringscriteria voor publieke investeringen: het uitwerken van een operationele theorie met een toepassing op de verkeersinfrastructuur. Leuven, Acco, 1975. 254 pp.

14. Niet toegekend.

15. VANOVERBEKE Lieven (1975)
Microeconomisch onderzoek van de sectoriële arbeidsmobiliteit. Leuven, Acco, 1975. 205 pp.

16. DAEMS Herman (1975)
The holding company: essays on financial intermediation, concentration and capital market imperfections in the Belgian economy. Leuven, K. U. Leuven, 1975. XII, 268 pp.

17. VAN ROMPUY Eric (1975)
Groot-Brittannië en de Europese monetaire integratie: een onderzoek naar de gevolgen van de Britse toetreding op de geplande Europese monetaire unie. Leuven, Acco, 1975. XIII, 222 pp.

18. MOESEN Wim (1975)
Het beheer van de staatsschuld en de termijnstructuur van de intrestvoeten met een toepassing voor België. Leuven, Vander, 1975. XVI, 250 pp.

19. LAMBRECHT Marc (1976)
Capacity constrained multi-facility dynamic lot-size problem. Leuven, K. U. Leuven, 1976. 165 pp.

20. RAYMAECKERS Erik (1976)
De mens in de onderneming en de theorie van het producenten-gedrag: een bijdrage tot transdisciplinaire analyse. Leuven, Acco, 1976. XIII, 538 pp.

21. TEJANO Albert (1976)
Econometric and input-output models in development planning: the case of the Philippines. Leuven, K. U. Leuven, 1976. XX, 297 pp.

22. MARTENS Bernard (1977)
Prijsbeleid en inflatie met een toepassing op België. Leuven, K. U. Leuven, 1977. IV, 253 pp.

23. VERHEIRSTRAETEN Albert (1977)
Geld, krediet en intrest in de Belgische financiële sector. Leuven, Acco, 1977. XXII, 377 pp.

24. GHEYSSENS Lieven (1977)
International diversification through the government bond market: a risk-return analysis. Leuven, s.n., 1977. 188 pp.

25. LEFEBVRE Chris (1977)
Boekhoudkundige verwerking en financiële verslaggeving van huurkooptransacties en verkopen op afbetaling bij ondernemingen die aan consumenten verkopen. Leuven, K. U. Leuven, 1977. 228 pp.

26. KESENNE Stefan (1978)
Tijdsallocatie en vrijetijdsbesteding: een econometrisch onderzoek. Leuven, s.n., 1978. 163 pp.

27. VAN HERCK Gustaaf (1978)
Aspecten van optimaal bedrijfsbeleid volgens het marktwaardecriterium: een risico-rendements-analyse. Leuven, K. U. Leuven, 1978. IV, 163 pp.

28. VAN POECK Andre (1979)
World price trends and price and wage development in Belgium: an investigation into the relevance of the Scandinavian model of inflation for Belgium. Leuven, s.n., 1979. XIV, 260 pp.

29. VOS Herman (1978)
De industriële technologieverwerving in Brazilië: een analyse. Leuven, s.n., 1978. onregelmatig gepagineerd.

30. DOMBRECHT Michel (1979)
Financial markets, employment and prices in open economies. Leuven, K. U. Leuven, 1979. 182 pp.

31. DE PRIL Nelson (1979)
Bijdrage tot de actuariële studie van het bonus-malussysteem. Brussel, OAB, 1979. 112 pp.

32. CARRIN Guy (1979)
Economic aspects of social security: a public economics approach. Leuven, K. U. Leuven, 1979. onregelmatig gepagineerd

33. REGIDOR Baldomero (1979)
An empirical investigation of the distribution of stock-market prices and weak-form efficiency of the Brussels stock exchange. Leuven, K. U. Leuven, 1979. 214 pp.

34. DE GROOT Roger (1979)
Ongelijkheden voor stop loss premies gebaseerd op E.T. systemen in het kader van de veralgemeende convexe analyse. Leuven, K. U. Leuven, 1979. 155 pp.

35. CEYSSENS Martin (1979)
On the peak load problem in the presence of rationizing by waiting. Leuven, K. U. Leuven, 1979. IX, 217 pp.

36. ABDUL RAZK Abdul (1979)
Mixed enterprise in Malaysia: the case study of joint venture between Malysian public corporations and foreign enterprises. Leuven, K. U. Leuven, 1979. 324 pp.

37. DE BRUYNE Guido (1980)
Coordination of economic policy: a game-theoretic approach. Leuven, K. U. Leuven, 1980. 106 pp.

38. KELLES Gerard (1980)
Demand, supply, price change and trading volume on financial markets of the matching-order type. Vraag, aanbod, koersontwikkeling en omzet op financiële markten van het Europese type. Leuven, K. U. Leuven, 1980. 222 pp.

39. VAN EECKHOUDT Marc (1980)
De invloed van de looptijd, de coupon en de verwachte inflatie op het opbrengstverloop van vastrentende finaciële activa. Leuven, K. U. Leuven, 1980. 294 pp.

40. SERCU Piet (1981)
Mean-variance asset pricing with deviations from purchasing power parity. Leuven, s.n., 1981. XIV, 273 pp.

41. DEQUAE Marie-Gemma (1981)
Inflatie, belastingsysteem en waarde van de onderneming. Leuven, K. U. Leuven, 1981. 436 pp.

42. BRENNAN John (1982)
An empirical investigation of Belgian price regulation by prior notifi-

cation: 1975 - 1979 - 1982. Leuven, K. U. Leuven, 1982. XIII, 386 pp.

43. COLLA Annie (1982)
Een econometrische analyse van ziekenhuiszorgen. Leuven, K. U. Leuven, 1982. 319 pp.

44. Niet toegekend.

45. SCHOKKAERT Eric (1982)
Modelling consumer preference formation. Leuven, K. U. Leuven, 1982. VIII, 287 pp.

46. DEGADT Jan (1982)
Specificatie van een econometrisch model voor vervuilingsproblemen met proeven van toepassing op de waterverontreiniging in België. Leuven, s.n., 1982. 2 V.

47. LANJONG Mohammad Nasir (1983)
A study of market efficiency and risk-return relationships in the Malaysian capital market. s.l., s.n., 1983. XVI, 287 pp.

48. PROOST Stef (1983)
De allocatie van lokale publieke goederen in een economie met een centrale overheid en lokale overheden. Leuven, s.n., 1983. onregelmatig gepagineerd.

49. VAN HULLE Cynthia (1983)
Shareholders' unanimity and optimal corporate decision making in imperfect capital markets. s.l., s.n., 1983. 147 pp. + appendix.

50. VAN WOUWE Martine (2/12/83)
Ordening van risico's met toepassing op de berekening van ultieme ruïnekansen. Leuven, s.n., 1983. 109 pp.

51. D'ALCANTARA Gonzague (15/12/83)
SERENA: a macroeconomic sectoral regional and national account econometric model for the Belgian economy. Leuven, K. U. Leuven, 1983. 595 pp.

52. D'HAVE Piet (24/02/84)
De vraag naar geld in België. Leuven, K. U. Leuven, 1984. XI, 318 pp.

53. MAES Ivo (16/03/84)
The contribution of J.R. Hicks to macro-economic and monetary theory. Leuven, K. U. Leuven, 1984. V, 224 pp.

54. SUBIANTO Bambang (13/09/84)
A study of the effects of specific taxes and subsidies on a firms' R&D investment plan. s.l., s.n., 1984. V, 284 pp.

55. SLEUWAEGEN Leo (26/10/84)
Location and investment decisions by multinational enterprises in Belgium and Europe. Leuven, K. U. Leuven, 1984. XII, 247 pp.

56. GEYSKENS Erik (27/03/85)
Produktietheorie en dualiteit. Leuven, s.n., 1985. VII, 392 pp.

57. COLE Frank (26/06/85)
Some algorithms for geometric programming. Leuven, K. U. Leuven, 1985. 166 pp.

58. STANDAERT Stan (26/09/86)
A study in the economics of repressed consumption. Leuven, K. U. Leuven, 1986. X, 380 pp.

59. DELBEKE Jos (03/11/86)
Trendperioden in de geldhoeveelheid van België 1877-1983: een theoretische en empirische analyse van de "Banking school" hypothese. Leuven, K. U. Leuven, 1986. XII, 430 pp.

60. VANTHIENEN Jan (08/12/86)
Automatiseringsaspecten van de specificatie, constructie en manipulatie van beslissingstabellen. Leuven, s.n., 1986. XIV, 378 pp.

61. LUYTEN Robert (30/04/87)
A systems-based approach for multi-echelon production/inventory systems. s.l., s.n., 1987. 3V.

62. MERCKEN Roger (27/04/87)
De invloed van de data base benadering op de interne controle. Leuven, s.n., 1987. XIII, 346 pp.

63. VAN CAYSEELE Patrick (20/05/87)
Regulation and international innovative activities in the pharmaceutical industry. s.l., s.n., 1987. XI, 169 pp.

64. FRANCOIS Pierre (21/09/87)
De empirische relevantie van de independence from irrelevant alternatives. Assumptie indiscrete keuzemodellen. Leuven, s.n., 1987. IX, 379 pp.

65. DECOSTER André (23/09/88)
Family size, welfare and public policy. Leuven, K. U. Leuven. Faculteit Economische en Toegepaste Economische Wetenschappen, 1988. XIII, 444 pp.

66. HEIJNEN Bart (09/09/88)
Risicowijziging onder invloed van vrijstellingen en herverzekeringen: een theoretische analyse van optimaliteit en premiebepaling. Leuven, K. U. Leuven. Faculteit Economische en Toegepaste Economische Wetenschappen, 1988. onregelmatig gepagineerd.

67. GEEROMS Hans (14/10/88)
Belastingvermijding. Theoretische analyse van de determinanten van de belastingontduiking en de belastingontwijking met empirische verificaties. Leuven, s.n., 1988. XIII, 409, 5 pp.

68. PUT Ferdi (19/12/88)
Introducing dynamic and temporal aspects in a conceptual (database) schema. Leuven, K. U. Leuven. Faculteit Economische en Toegepaste Economische Wetenschappen, 1988. XVIII, 415 pp.

69. VAN ROMPUY Guido (13/01/89)
A supply-side approach to tax reform programs. Theory and empirical evidence for Belgium. Leuven, K. U. Leuven. Faculteit Economische en Toegepaste Economische Wetenschappen, 1989. XVI, 189, 6 pp.

70. PEETERS Ludo (19/06/89)
Een ruimtelijk evenwichtsmodel van de graanmarkten in de E.G.: empirische specificatie en beleidstoepassingen. Leuven, K. U. Leuven. Faculteit Economische en Toegepaste Economische Wetenschappen, 1989. XVI, 412 pp.

71. PACOLET Jozef (10/11/89)
Marktstructuur en operationele efficiëntie in de Belgische financiële sector. Leuven, K. U. Leuven. Faculteit Economische en Toegepaste Economische Wetenschappen, 1989. XXII, 547 pp.

72. VANDEBROEK Martina (13/12/89)
Optimalisatie van verzekeringscontracten en premieberekeningsprincipes. Leuven, K. U. Leuven. Faculteit Economische en Toegepaste Economische Wetenschappen, 1989. 95 pp.

73. WILLEKENS Francois (1990)
Determinance of government growth in industrialized countries with applications to Belgium. Leuven, K. U. Leuven. Faculteit Economische en Toegepaste Economische Wetenschappen, 1990. VI, 332 pp.

74. VEUGELERS Reinhilde (02/04/90)
Scope decisions of multinational enterprises. Leuven, K. U. Leuven. Faculteit Economische en Toegepaste Economische Wetenschappen, 1990. V, 221 pp.

75. KESTELOOT Katrien (18/06/90)
Essays on performance diagnosis and tacit cooperation in international oligopolies. Leuven, K. U. Leuven. Faculteit Economische en Toegepaste Economische Wetenschappen, 1990. 227 pp.

76. WU Changqi (23/10/90) Strategic aspects of oligopolistic vertical integration. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1990. VIII, 222 pp.

77. ZHANG Zhaoyong (08/07/91)
A disequilibrium model of China's foreign trade behaviour. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1991. XII, 256 pp.

78. DHAENE Jan (25/11/91)
Verdelingsfuncties, benaderingen en foutengrenzen van stochastische grootheden geassocieerd aan verzekeringspolissen en -portefeuilles. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1991. 146 pp.

79. BAUWELINCKX Thierry (07/01/92)
Hierarchical credibility techniques. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1992. 130 pp.

80. DEMEULEMEESTER Erik (23/3/92)
Optimal algorithms for various classes of multiple resource-constrained project scheduling problems. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1992. 180 pp.

81. STEENACKERS Anna (1/10/92)
Risk analysis with the classical actuarial risk model: theoretical extensions and applications to Reinsurance. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1992. 139 pp.

82. COCKX Bart (24/09/92)
The minimum income guarantee. Some views from a dynamic perspective. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1992. XVII, 401 pp.

83. MEYERMANS Eric (06/11/92)
Econometric allocation systems for the foreign exchange market: Specification, estimation and testing of transmission mechanisms under currency substitution. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1992. XVIII, 343 pp.

84. CHEN Guoqing (04/12/92)
Design of fuzzy relational databases based on fuzzy functional dependency. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1992. 176 pp.

85. CLAEYS Christel (18/02/93)
Vertical and horizontal category structures in consumer decision making: The nature of product hierarchies and the effect of brand typicality. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1993. 348 pp.

86. CHEN Shaoxiang (25/03/93)
The optimal monitoring policies for some stochastic and dynamic production processes. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1993. 170 pp.

87. OVERWEG Dirk (23/04/93)
Approximate parametric analysis and study of cost capacity management of computer configurations. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1993. 270 pp.

88. DEWACHTER Hans (22/06/93)
Nonlinearities in speculative prices: The existence and persistence of nonlinearity in foreign exchange rates. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1993. 151 pp.

89. LIN Liangqi (05/07/93)
Economic determinants of voluntary accounting choices for R & D expenditures in Belgium. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1993. 192 pp.

90. DHAENE Geert (09/07/93)
Encompassing: formulation, properties and testing. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1993. 117 pp.

91. LAGAE Wim (20/09/93)
Marktconforme verlichting van soevereine buitenlandse schuld door private crediteuren: een neo-institutionele analyse. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1993. 241 pp.

92. VAN DE GAER Dirk (27/09/93)
Equality of opportunity and investment in human capital. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1993. 172 pp.

93. SCHROYEN Alfred (28/02/94)
Essays on redistributive taxation when monitoring is costly. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1994. 203 pp. + V.

94. STEURS Geert (15/07/94)
Spillovers and cooperation in research and development. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1994. 266 pp.

95. BARAS Johan (15/09/94)
The small sample distribution of the Wald, Lagrange multiplier and likelihood ratio tests for homogeneity and symmetry in demand analysis: a Monte Carlo study. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1994. 169 pp.

96. GAEREMYNCK Ann (08/09/94)
The use of depreciation in accounting as a signalling device. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1994. 232 pp.

97. BETTENDORF Leon (22/09/94)
A dynamic applied general equilibrium model for a small open economy. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1994. 149 pp.

98. TEUNEN Marleen (10/11/94)
Evaluation of interest randomness in actuarial quantities. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1994. 214 pp.

99. VAN OOTEGEM Luc (17/01/95)
An economic theory of private donations. Leuven. K. U. Leuven, Fa-

culteit Economische en Toegepaste Economische Wetenschappen, 1995. 236 pp.

100. DE SCHEPPER Ann (20/03/95)
Stochastic interest rates and the probabilistic behaviour of actuarial functions. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1995. 211 pp.

101. LAUWERS Luc (13/06/95)
Social choice with infinite populations. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1995. 79 pp.

102. WU Guang (27/06/95)
A systematic approach to object-oriented business modeling. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1995. 248 pp.

103. WU Xueping (21/08/95)
Term structures in the Belgian market: model estimation and pricing error analysis. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1995. 133 pp.

104. PEPERMANS Guido (30/08/95)
Four essays on retirement from the labor force. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1995. 128 pp.

105. ALGOED Koen (11/09/95)
Essays on insurance: a view from a dynamic perspective. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1995. 136 pp.

106. DEGRYSE Hans (10/10/95)
Essays on financial intermediation, product differentiation, and market structure. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1995. 218 pp.

107. MEIR Jos (05/12/95)
Het strategisch groepsconcept toegepast op de Belgische financiële sec-

tor. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1995. 257 pp.

108. WIJAYA Miryam Lilian (08/01/96)
Voluntary reciprocity as an informal social insurance mechanism: a game theoretic approach. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1996. 124 pp.

109. VANDAELE Nico (12/02/96)
The impact of lot sizing on queueing delays: multi-product, multi-machine models. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1996. 243 pp.

110. GIELENS Geert (27/02/96)
Some essays on discrete time target zones and their tails. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1996. 131 pp.

111. GUILLAUME Dominique (20/03/96)
Chaos, randomness and order in the foreign exchange markets. Essays on the modelling of the markets. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1996. 171 pp.

112. DEWIT Gerda (03/06/96)
Essays on export insurance subsidization. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1996. 186 pp.

113. VAN DEN ACKER Carine (08/07/96)
Belief-function theory and its application to the modeling of uncertainty in financial statement auditing. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1996. 147 pp.

114. IMAM Mahmood Osman (31/07/96)
Choice of IPO Flotation Methods in Belgium in an Asymmetric Information Framework and Pricing of IPO's in the Long-Run. Leuven,

K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1996. 221 pp.

115. NICAISE Ides (06/09/96)
Poverty and Human Capital. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1996. 209 pp.

116. EYCKMANS Johan (18/09/97)
On the Incentives of Nations to Join International Environmental Agreements. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1997. XV + 348 pp.

117. CRISOLOGO-MENDOZA Lorelei (16/10/97)
Essays on Decision Making in Rural Households: a study of three villages in the Cordillera Region of the Philippines. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1997. 256 pp.

118. DE REYCK Bert (26/01/98)
Scheduling Projects with Generalized Precedence Relations: Exact and Heuristic Procedures. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1998. XXIV + 337 pp.

119. VANDEMAELE Sigrid (30/04/98)
Determinants of Issue Procedure Choice within the Context of the French IPO Market: Analysis within an Asymmetric Information Framework. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1998. 241 pp.

120. VERGAUWEN Filip (30/04/98)
Firm Efficiency and Compensation Schemes for the Management of Innovative Activities and Knowledge Transfers. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1998. VIII + 175 pp.

121. LEEMANS Herlinde (29/05/98)
The Two-Class Two-Server Queueing Model with Nonpreemptive Heterogeneous Priority Structures. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1998. 211 pp.

122. GEYSKENS Inge (4/09/98)
Trust, Satisfaction, and Equity in Marketing Channel Relationships. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1998. 202 pp.

123. SWEENEY John (19/10/98)
Why Hold a Job? The Labour Market Choice of the Low-Skilled. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1998. 278 pp.

124. GOEDHUYS Micheline (17/03/99)
Industrial Organisation in Developing Countries, Evidence from Côte d'Ivoire. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1999. 251 pp.

125. POELS Geert (16/04/99)
On the Formal Aspects of the Measurement of Object-Oriented Software Specifications. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1999. 507 pp.

126. MAYERES Inge (25/05/99)
The Control of Transport Externalities: A General Equilibrium Analysis. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1999. XIV + 294 pp.

127. LEMAHIEU Wilfried (5/07/99)
Improved Navigation and Maintenance through an Object-Oriented Approach to Hypermedia Modelling. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1999. 284 pp.

128. VAN PUYENBROECK Tom (8/07/99)
Informational Aspects of Fiscal Federalism. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1999. 192 pp.

129. VAN DEN POEL Dirk (5/08/99)
Response Modeling for Database Marketing Using Binary Classifica-

tion. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1999. 342 pp.

130. GIELENS Katrijn (27/08/99)
International Entry Decisions in the Retailing Industry: Antecedents and Performance Consequences. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1999. 336 pp.

131. PEETERS Anneleen (16/12/99)
Labour Turnover Costs, Employment and Temporary Work. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1999. 207 pp.

132. VANHOENACKER Jurgen (17/12/99)
Formalizing a Knowledge Management Architecture Meta-Model for Integrated Business Process Management. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 1999. 252 pp.

133. NUNES Paulo (20/03/2000)
Contingent Valuation of the Benefits of Natural Areas and its Warmglow Component. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2000. XXI + 282 pp.

134. VAN DEN CRUYCE Bart (7/04/2000)
Statistische discriminatie van allochtonen op jobmarkten met rigide lonen. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2000. XXIII + 441 pp.

135. REPKINE Alexandre (15/03/2000)
Industrial restructuring in countries of Central and Eastern Europe: Combining branch-, firm- and product-level data for a better understanding of Enterprises' behaviour during transition towards market economy. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2000. VI + 147 pp.

136. AKSOY Yunus (21/06/2000)
Essays on international price rigidities and exchange rates. Leuven,

K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2000. IX + 236 pp.

137. RIYANTO Yohanes Eko (22/06/2000)
Essays on the internal and external delegation of authority in firms. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2000. VIII + 280 pp.

138. HUYGHEBAERT Nancy (20/12/2000)
The Capital Structure of Business Start-ups. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2000. VIII + 332 pp.

139. FRANCKX Laurent (22/01/2001)
Ambient Inspections and Commitment in Environmental Enforcement. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2001. VIII + 286 pp.

140. VANDILLE Guy (16/02/2001)
Essays on the Impact of Income Redistribution on Trade. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2001. VIII + 176 pp.

141. MARQUERING Wessel (27/04/2001)
Modeling and Forecasting Stock Market Returns and Volatility. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2001. V + 267 pp.

142. FAGGIO Giulia (07/05/2001)
Labor Market Adjustment and Enterprise Behavior in Transition. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2001. 150 pp.

143. GOOS Peter (30/05/2001)
The Optimal Design of Blocked and Split-plot experiments. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2001. X + 224 pp.

144. LABRO Eva (01/06/2001)
Total Cost of Ownership Supplier Selection based on Activity Based Costing and Mathematical Programming. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2001. 217 pp.

145. VANHOUCKE Mario (07/06/2001)
Exact Algorithms for various Types of Project Scheduling Problems. Nonregular Objectives and time/cost Trade-offs. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2001. 316 pp.

146. BILSEN Valentijn (28/08/2001)
Entrepreneurship and Private Sector Development in Central European Transition Countries. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2001. XVI + 188 pp.

147. NIJS Vincent (10/08/2001)
Essays on the dynamic Category-level Impact of Price promotions. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2001.

148. CHERCHYE Laurens (24/09/2001)
Topics in Non-parametric Production and Efficiency Analysis. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2001. VII + 169 pp.

149. VAN DENDER Kurt (15/10/2001)
Aspects of Congestion Pricing for Urban Transport. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2001. VII + 203 pp.

150. CAPEAU Bart (26/10/2001)
In defence of the excess demand approach to poor peasants' economic behaviour. Theory and Empirics of non-recursive agricultural household modelling. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2001. XIII + 286 pp.

151. CALTHROP Edward (09/11/2001)
Essays in urban transport economics. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2001.

152. VANDER BAUWHEDE Heidi (03/12/2001)
Earnings management in an Non-Anglo-Saxon environment. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2001. 408 pp.

153. DE BACKER Koenraad (22/01/2002)
Multinational firms and industry dynamics in host countries: the case of Belgium. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2002. VII + 165 pp.

154. BOUWEN Jan (08/02/2002)
Transactive memory in operational workgroups. Concept elaboration and case study. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2002. 319 pp. + appendix 102 pp.

155. VAN DEN BRANDE Inge (13/03/2002)
The psychological contract between employer and employee: a survey among Flemish employees. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2002. VIII + 470 pp.

156. VEESTRAETEN Dirk (19/04/2002)
Asset Price Dynamics under Announced Policy Switching. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2002. 176 pp.

157. PEETERS Marc (16/05/2002)
One Dimensional Cutting and Packing: New Problems and Algorithms. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2002. IX + 247 pp.

158. SKUDELNY Frauke (21/05/2002)
Essays on The Economic Consequences of the European Monetary

Union. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2002.

159. DE WEERDT Joachim (07/06/2002)
Social Networks, Transfers and Insurance in Developing countries. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2002. VI + 129 pp.

160. TACK Lieven (25/06/2002)
Optimal Run Orders in Design of Experiments. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2002. XXXI + 344 pp.

161. POELMANS Stephan (10/07/2002)
Making Workflow Systems work. An investigation into the Importance of Task-appropriation fit, End-user Support and other Technological Characteristics. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2002. 237 pp.

162. JANS Raf (26/09/2002)
Capacitated Lot Sizing Problems: New Applications, Formulations and Algorithms. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2002.

163. VIAENE Stijn (25/10/2002)
Learning to Detect Fraud from enriched Insurance Claims Data (Context, Theory and Applications). Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2002. 315 pp.

164. AYALEW Tekabe (08/11/2002)
Inequality and Capital Investment in a Subsistence Economy. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2002. V + 148 pp.

165. MUES Christophe (12/11/2002)
On the Use of Decision Tables and Diagrams in Knowledge Modeling and Verification. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2002. 222 pp.

166. BROCK Ellen (13/03/2003)
The Impact of International Trade on European Labour Markets. K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2002.

167. VERMEULEN Frederic (29/11/2002)
Essays on the collective Approach to Household Labour Supply. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2002. XIV + 203 pp.

168. CLUDTS Stephan (11/12/2002)
Combining participation in decision-making with financial participation: theoretical and empirical perspectives. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2002. XIV + 247 pp.

169. WARZYNSKI Frederic (09/01/2003)
The dynamic effect of competition on price cost margins and innovation. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2003.

170. VERWIMP Philip (14/01/2003)
Development and genocide in Rwanda: a political economy analysis of peasants and power under the Habyarimana regime. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2003.

171. BIGANO Andrea (25/02/2003)
Environmental regulation of the electricity sector in a European Market Framework. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2003. XX + 310 pp.

172. MAES Konstantijn (24/03/2003)
Modeling the Term Structure of Interest Rates Across Countries. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2003. V+246 pp.

173. VINAIMONT Tom (26/02/2003)
The performance of One- versus Two-Factor Models of the Term Struc-

ture of Interest Rates. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen. 2003.

174. OOGHE Erwin (15/04/2003)
Essays in multi-dimensional social choice. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2003. VIII+108 pp.

175. FORRIER Anneleen (25/04/2003)
Temporary employment, employability and training. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2003.

176. CARDINAELS Eddy (28/04/2003)
The role of cost system accuracy in managerial decision making. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2003. 144 pp.

177. DE GOEIJ Peter (02/07/2003)
Modeling Time-Varying Volatility and Interest Rates. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2003. VII+225 pp.

178. LEUS Roel (19/09/2003)
The generation of stable project plans. Complexity and exact algorithms. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2003.

179. MARINHEIRO Carlos (23/09/2003)
EMU and fiscal stabilisation policy: the case of small countries. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2003.

180. BAESSENS Bart (24/09/2003)
Developing intelligent systems for credit scoring using machine learning techniques. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2003.

181. KOCZY Laszlo (18/09/2003)
Solution concepts and outsider behaviour in coalition formation games. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2003.

182. ALTOMONTE Carlo (25/09/2003)
Essays on Foreign Direct Investment in transition countries: learning from the evidence. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2003.

183. DRIES Liesbeth (10/11/2003)
Transition, Globalisation and Sectoral Restructuring: Theory and Evidence from the Polish Agri-Food Sector. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2003.

184. DEVOOGHT Kurt (18/11/2003)
Essays On Responsibility-Sensitive Egalitarianism and the Measurement of Income Inequality. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2003.

185. DELEERSNYDER Barbara (28/11/2003)
Marketing in Turbulent Times. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2003.

186. ALI Daniel (19/12/2003)
Essays on Household Consumption and Production Decisions under Uncertainty in Rural Ethiopia. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2003.

187. WILLEMS Bert (14/01/2004)
Electricity networks and generation market power. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2004.

188. JANSSENS Gust (30/01/2004)
Advanced Modelling of Conditional Volatility and Correlation in Financial Markets. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2004.

189. THOEN Vincent (19/01/2004)
On the valuation and disclosure practices implemented by venture capital providers. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2004.

190. MARTENS Jurgen (16/02/2004)
A fuzzy set and stochastic system theoretic technique to validate simulation models. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2004.

191. ALTAVILLA Carlo (21/05/2004)
Monetary policy implementation and transmission mechanisms in the Euro area. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2004.

192. DE BRUYNE Karolien (07/06/2004)
Essays in the location of economic activity. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2004.

193. ADEM Jan (25/06/2004)
Mathematical programming approaches for the supervised classification problem. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2004.

194. LEROUGE Davy (08/07/2004)
Predicting Product Preferences: the effect of internal and external cues. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2004.

195. VANDENBROECK Katleen (16/07/2004)
Essays on output growth, social learning and land allocation in agriculture: micro-evidence from Ethiopia and Tanzania. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2004.

196. GRIMALDI Maria (03/09/2004)
The exchange rate, heterogeneity of agents and bounded rationality. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2004.

197. SMEDTS Kristien (26/10/2004)
Financial integration in EMU in the framework of the no-arbitrage theory. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2004.

198. KOEVOETS Wim (12/11/2004)
Essays on Unions, Wages and Employment. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2004.

199. CALLENS Marc (22/11/2004)
Essays on multilevel logistic Regression. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2004.

200. RUGGOO Arvind (13/12/2004)
Two stage designs robust to model uncertainty. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2004.

201. HOORELBEKE Dirk (28/01/2005)
Bootstrap and Pivoting Techniques for Testing Multiple Hypotheses. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2005.

202. ROUSSEAU Sandra (17/02/2005)
Selecting Environmental Policy Instruments in the Presence of Incomplete Compiance. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2005.

203. VAN DER MEULEN Sofie (17/02/2005)
Quality of Financial Statements: Impact of the external auditor and applied accounting standards. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2005.

204. DIMOVA Ralitza (21/02/2005)
Winners and Losers during Structural Reform and Crisis: the Bulgarian Labour Market Perspective. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2005.

205. DARKIEWICZ Grzegorz (28/02/2005)
Value-at-risk in Insurance and Finance: the Comonotonicity Approach. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2005.

206. DE MOOR Lieven (20/05/2005)
The Structure of International Stock Returns: Size, Country and Sector Effects in Capital Asset Pricing. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2005.

207. EVERAERT Greetje (27/06/2005)
Soft Budget Constraints and Trade Policies: The Role of Institutional and External Constraints. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2005.

208. SIMON Steven (06/07/2005)
The Modeling and Valuation of complex Derivatives: The Impact of the Choice of the term structure model. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2005.

209. MOONEN Linda (23/09/2005)
Algorithms for some Graph-Theoretical Optimization Problems. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2005.

210. COUCKE Kristien (21/09/2005)
Firm and industry adjustment under de-industrialisation and globalization of the Belgian economy. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2005.

211. DECAMPS Marc (21/10/2005)
Some actuarial and financial applications of generalized diffusion processes. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2005.

212. KIM Helena (29/11/2005)
Escalation games: an instrument to analyze conflicts. The strategic approach to the bargaining problem. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2005.

213. GERMENJI Etleva (06/01/2006)
Essays on the Economics of Emigration from Albania. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

214. BELIEN Jeroen (18/01/2006)
Exact and Heuristic Methodologies for Scheduling in Hospitals: Problems, Formulations and Algorithms. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

215. JOOSSENS Kristel (10/02/2006)
Robust discriminant analysis. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

216. VRANKEN Liesbet (13/02/2006)
Land markets and production efficiency in transition economies. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

217. VANSTEENKISTE Isabel (22/02/2006)
Essays on non-linear modelling in international macroeconomics. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

218. WUYTS Gunther (31/03/2006)
Essays on the liquidity of financial markets. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

219. DE BLANDER Rembert (28/04/2006)
Essays on endogeneity and parameter heterogeneity in cross-section and panel data. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

220. DE LOECKER Jan (12/05/2006)
Industry dynamics and productivity. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

221. LEMMENS Aurélie (12/05/2006)
Advanced classification and time-series methods in marketing. Leu-

ven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

222. VERPOORTEN Marijke (22/05/2006)
Conflict and survival: an analysis of shocks, coping strategies and economic mobility in Rwanda, 1990-2002. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

223. BOSMANS Kristof (26/05/2006)
Measuring economic inequality and inequality aversion. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

224. BRENKERS Randy (29/05/2006)
Policy reform in a market with differentiated products: applications from the car market. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

225. BRUYNEEL Sabrina (02/06/2006)
Self-control depletion: Mechanisms and its effects on consumer behavior. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

226. FAEMS Dries (09/06/2006)
Collaboration for innovation: Processes of governance and learning. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

227. BRIERS Barbara (28/06/2006)
Countering the scrooge in each of us: on the marketing of cooperative behavior. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

228. ZANONI Patrizia (04/07/2006)
Beyond demography: Essays on diversity in organizations. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

229. VAN DEN ABBEELE Alexandra (11/09/2006)
Management control of interfirm relations: the role of information. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

230. DEWAELHEYNS Nico (18/09/2006)
Essays on internal capital markets, bankruptcy and bankruptcy reform. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

231. RINALDI Laura (19/09/2006)
Essays on card payments and household debt. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

232. DUTORDOIR Marie (22/09/2006)
Determinants and stock price effects of Western European convertible debt offerings: an empirical analysis. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

233. LYKOGIANNI Elissavet (20/09/2006)
Essays on strategic decisions of multinational enterprises: R&D decentralization, technology transfers and modes of foreign entry. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

234. ZOU Jianglei (03/10/2006)
Inter-firm ties, plant networks, and multinational firms: essays on FDI and trade by Japanse firms. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

235. GEYSKENS Kelly (12/10/2006)
The ironic effects of food temptations on self-control performance. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

236. BRUYNSEELS Liesbeth (17/10/2006)
Client strategic actions, going-concern audit opinions and audit report-

ing errors. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.

237. KESSELS Roselinde (23/10/2006)
Optimal designs for the measurement of consumer preferences. Leuven, K. U. Leuven, Faculteit Economische en Toegepaste Economische Wetenschappen, 2006.