The Design of Efficient Stated Choice Experiments

by

Stephen A. BUSH

Submitted to the Department of Mathematical Sciences, Faculty of Science
in partial fulfilment of the requirements for the degree of

Doctor of Philosophy in Mathematics

at the

UNIVERSITY OF TECHNOLOGY, SYDNEY.

February, 2010

CERTIFICATE OF AUTHORSHIP/ORIGINALITY

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature of Candidate

____

Acknowledgements

I would first like to thank my supervisor, Professor Deborah Street, for all of her constructive feedback, patience, wise advice, and for her good humour throughout the whole degree. I would also like to thank Dr Leonie Burgess for all of her feedback, and for writing the original routines for checking designs and running simulations, which I have since modified for the models covered in this thesis, and Dr Narelle Smith for all of the sound advice and encouragement that she provided. In addition, I would also like to thank those staff at the Department of Mathematical Sciences at UTS not previously mentioned, who have also regularly enquired on my progress and provided the right words at the right time. I feel privileged to work with such a great group of people.

I would like to acknowledge the Centre for the Study of Choice (CenSoC) for providing the doctoral scholarship that allowed me to focus on completing this thesis in a timely manner.

Last but certainly not least I would like to thank my family, and in particular my partner Sam, for all of their support, patience and understanding. I consider myself very fortunate to have such a great family.

Contents

List of Tables

List of Figures

List of Notation

α	the primitive root in $GF[\ell]$ (Chapter 6 only)
$\alpha_{\boldsymbol{e},a,b}$	the number of times $\boldsymbol{e} = (e_1, e_2, \ldots, e_k)$ appears as a difference between the items in positions a and b of the choice set
$\boldsymbol{\beta}$	a vector containing the effects of each level of each attribute, and the effects of each combination of levels
$\delta_{T_i \text{ in pos } a}$	an indicator variable which equals 1 if item T_i appears in position a of the choice set
η	the utility threshold for equality of preferences
ϵ_i	the stochastic component of the utility of the item T_i
$\boldsymbol{\gamma}$	a vector containing the log merits for each item
$\lambda_{\{i,j\}}$	the proportion of choice sets that are $\{T_i, T_j\}$
$\lambda_{(i,j)}$	the proportion of choice sets that are (T_i, T_j)
λ_C	the proportion of the choice sets that are the choice set C
$\lambda_{T_i \text{ in pos } a}$	the proportion of choice sets with item T_i in position a of the choice set
$\lambda_{T_i \text{ in pos } a, T_j \text{ in pos } b}$	the proportion of choice sets with item T_i in position a of the choice set and T_j in position b of the choice set
$\lambda_{\text{att } q=x \text{ in pos } a}$	the proportion of choice sets where the item in the item in position a of the choice set has the q^{th} attribute at the x^{th} level
$\lambda_{\text{att } q_1=x_1, q_2=x_2 \text{ in pos } a}$	the proportion of choice sets where the item in the item in position a of the choice set has the q_1^{th} attribute at the x_1^{th} level and the q_2^{th} attribute at the x_2^{th} level
$\Lambda(\boldsymbol{\pi})$	the information matrix for the estimation of the entries in $\boldsymbol{\gamma}$
$\Lambda(\boldsymbol{\pi}_0)_{\text{B–T}}$	the information matrix for the estimation of contrasts in $\boldsymbol{\gamma}$ when the Bradley–Terry model is used
$\Lambda(\boldsymbol{\pi}_0)_{\text{MNL}}$	the information matrix for the estimation of contrasts in $\boldsymbol{\gamma}$ when the MNL model is used

$\Lambda(\boldsymbol{\pi}_0, \nu)_{\text{DAV}}$	the information matrix for the estimation of the entries in $\boldsymbol{\gamma}$ and ν when the (generalised) Davidson ties model is used
$\Lambda(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\text{D–B}}$	the information matrix for the estimation of contrasts in $\boldsymbol{\gamma}$ and $\boldsymbol{\psi}$ when the (generalised) Davidson–Beaver position effects model is used
ν	the ties parameter for the Davidson ties model and the generalised Davidson ties model
π_i	the merit of item T_i
π_{i_b}	the merit of the item in position b of the choice set
$\boldsymbol{\pi}$	a vector containing the merits of each item
$\boldsymbol{\pi}_0$	the vector of π_is under the null hypothesis of equal selection probabilities
ψ_a	the position effect parameter for position a of the choice set
ψ_L	the linear component of the position effect
ψ_Q	the quadratic component of the position effect
$\boldsymbol{\psi}$	the vector containing $\psi_1, \ldots, \psi_m$
$\Sigma(\boldsymbol{\pi})_a$	the variance–covariance matrix for $\sqrt{sN}\boldsymbol{\gamma}$
$\Sigma(\boldsymbol{\pi})$	the variance–covariance matrix for the parameters of interest
θ	the ties parameter in the Rao–Kupper ties model
τ_i	the main effect of the i^{th} attribute, which takes 2 levels
$\tau_{i,\text{LIN}}$	the linear component of the main effect of the i^{th} attribute
ξ	a design
$\mathfrak{X}$	the class of competing designs
$\otimes$	Kronecker product
BlkDiag[]	a block diagonal matrix
$\binom{m}{x}$	the number of ways of choosing x objects from a set of m distinct objects without repetitions
$\bar{A}$	the set of items that do not appear in the choice experiment
$A_{\text{eff}}(\xi)$	the A–efficiency of a design $\xi \in \mathfrak{X}$ over the class of competing designs
$a_{k,i}$	the frequency of choice sets that differ in i attributes
B	a contrast matrix
B_a	a contrast matrix containing those effects that we assume are negligible
$B_{\bar{A}}$	the contrast coefficients that correspond to those items that do not appear in the choice experiment

$B_{F(q)}$	the rows of a contrast matrix corresponding to the main effects of attribute q
$B_{F,1}$	the rows of the contrast matrix for the items in F that correspond to the main effects of the first $k-p$ attributes
$B_{F,2}$	the rows of the contrast matrix for the items in F that correspond to the main effects of the last p attributes
B_h	a contrast matrix containing those effects that we are interested in estimating
B_ℓ	a contrast matrix containing the orthogonal polynomial contrast coefficients for an ℓ level attribute
B_M	a contrast matrix containing the contrast coefficients for the attribute main effects
B_T	a contrast matrix containing the contrast coefficients for the two–factor interactions between attributes
$\boldsymbol{b}_q$	the column of an orthogonal array corresponding to the q^{th} attribute
C	the information matrix for the estimation of the effects of interest
$C = \{T_{i_1}, \ldots, T_{i_m}\}$	the unordered choice set containing the items $T_{i_1}, \ldots, T_{i_m}$
$C = (T_{i_1}, \ldots, T_{i_m})$	the ordered choice set containing the items $T_{i_1}, \ldots, T_{i_m}$
$C(\boldsymbol{\pi}_0)_{\text{B–T}}$	the information matrix for the estimation of contrasts in $B\boldsymbol{\beta}$ when the Bradley–Terry model is used
$C(\boldsymbol{\pi}_0)_{\text{MNL}}$	the information matrix for the estimation of contrasts in $B\boldsymbol{\beta}$ when the MNL model is used
$C(\boldsymbol{\pi}_0, \nu)_{\text{DAV}}$	the information matrix for the estimation of contrasts in $B\boldsymbol{\beta}$ and ν when the (generalised) Davidson ties model is used
$C(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\text{D–B}}$	the information matrix for the estimation of contrasts in $B\boldsymbol{\beta}$ and $B_\psi\boldsymbol{\psi}$ when the (generalised) Davidson–Beaver position effects model is used
$C(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{MP}$	the information matrix for the estimation of main effects and position effects
$(C(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{MP})_{CLS}$	the information matrix for the estimation of main effects and position effects when a complete Latin square based design is used
$(C(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{MP})_{S-B}$	the information matrix for the estimation of main effects and position effects when a design from Burgess and Street [2005] is used
$C(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{MTP}$	the information matrix for the estimation of main effects plus two–factor interactions and position effects

$c_{\boldsymbol{v}_j}$	the number of choice sets containing the item $00\ldots0$ with difference vector $\boldsymbol{v}_j$
$c_{\boldsymbol{v}_j,a}$	the number of ordered choice sets with ordered difference vector $\boldsymbol{v}_j$ that contain the item $00\ldots0$ in position a of the choice set
$D_{\boldsymbol{d}}$	a $t \times t$ matrix with entries either 0 or 1 such that there is a 1 is position (i,j) if the items T_i and T_j have difference $\boldsymbol{d}$
$D_{\text{eff}}(\xi)$	the D–efficiency of a design $\xi \in \mathfrak{X}$ over the class of competing designs
d_{ij}	the number of attributes that differ between items T_i and T_j
$\boldsymbol{d}$	a difference
F	a starting design
$f(\boldsymbol{x},\boldsymbol{\theta})$	the joint probability distribution function of the independent sample $\boldsymbol{X} = (x_1,\ldots,X_n)$
G	a set of m generators
$G()$	the Lagrangian function for obtaining the maximum likelihood estimators
$\boldsymbol{g}$	a generator
I_ℓ	an $\ell \times \ell$ identity matrix
$I(\boldsymbol{\pi})$	the information matrix for the estimation of the entries in $\boldsymbol{\pi}$
i_q	an indicator that equals 1 if the q^{th} position of the difference $\boldsymbol{d}$ equals 1
$i_{\boldsymbol{v}_j}$	an indicator of whether all choice sets with difference vector $\boldsymbol{v}_j$ appear in the experiment
$I_\xi(\boldsymbol{\theta})$	the Fisher information matrix for the estimation of the entries in $\boldsymbol{\theta}$
J_ℓ	an $\ell \times \ell$ matrix of 1s
$\boldsymbol{j}_\ell$	a vector containing ℓ 1s
k	the number of attributes describing an item
ℓ_i	the number of levels that the i^{th} attribute may take
$L_{i,j}$	the $(i,j)^{\text{th}}$ entry in a Latin square
$L(\boldsymbol{x},\boldsymbol{\theta})$	the likelihood function
m	the number of options in each choice set
m_{ijk}	the expected number of times the k^{th} outcome will occur when the choice set $\{T_i,T_j\}$ or (T_i,T_j) are presented
N	the number of choice sets in a choice experiment

$n_{\{i,j\}}$	the number of times that the choice set $\{T_i, T_j\}$ appears in the experiment
$n_{(i,j)}$	the number of times that the choice set (T_i, T_j) appears in the experiment
n_C	the number of times that the choice set C appears in the experiment
p	the number of attribute contrasts being estimated
P_i	an $\ell \times \ell$ permutation matrix where $(P_i)_{x,y} = 1$ if and only if $x + i = y \mod \ell$
P_{ℓ_q,e_q}	a $t \times t$ matrix with entries either 0 or 1 such that there is a 1 is position (t_1, t_2) if $t_2 - t_1 = e_q$
Q_i	an $\ell \times \ell$ permutation matrix where $(Q_i)_{x,y} = 1$ if and only if $x + i = y$ in $GF[\ell]$
S_q	the optimal number of non–zero entries in a difference vector corresponding to a particular q level attribute
T_i	an item
T_{i_b}	the item in position b of the choice set C
U_i	the utility of the item T_i
V_i	the deterministic component of utility of the item T_i
$\boldsymbol{v}$	a difference vector
$w_{i\|C,\alpha}$	an indicator variable that equals 1 if respondent α selects item T_i when presented with the choice set C
$w_{\{i,j\}\|C,\alpha}$	an indicator variable that equals 1 if respondent α selects finds items T_i and T_j equally attractive when presented with the choice set C
$\boldsymbol{w}$	a vector containing all of the selection indicators for an experiment
$x_{\boldsymbol{v}_j;\boldsymbol{d}}$	the number of times the difference $\boldsymbol{d}$ appears in the difference vector $\boldsymbol{v}_j$
$x_{\boldsymbol{v}_j;\boldsymbol{d},a,b}$	the number of times that the difference $\boldsymbol{d}$ appears as a difference between positions a and b of the ordered difference vector $\boldsymbol{v}_j$
y_i	the proportion of pairs in a choice experiment that differ in the levels of i attributes
$y_{\boldsymbol{d}}$	the proportion of pairs of items in the choice experiment with difference $\boldsymbol{d}$
$y_{\boldsymbol{d},a,b}$	the proportion of pairs in the choice experiment that are in positions a and b of the choice set and have difference vector $\boldsymbol{d}$

Abstract

Making choices is a fundamental part of life. Whether it be the food that we eat, how we get from A to B, or the things that we do or do not purchase, choices are made all of the time. The ability to understand and influence these choices is valuable in many areas such as marketing, health economics, tourism, transportation research, and public policy. Choice experiments allow researchers in these areas to show respondents sets of options, described by attributes, and use the attributes of the chosen options to determine how important each of the attributes are to the 'attractiveness' of any option. From this information market share or policy acceptability can be predicted.

In this thesis we look at optimal designs for the multinomial logit (MNL) model, and for two extensions of this model. The first extension incorporates tied preferences, and is based on the extension of the Bradley–Terry model introduced by Davidson [1970]. The second extension allows the researcher to estimate the effect that the position of an item in the set of alternatives has on the perceived merit of the item. This extension is based on the extension of the Bradley–Terry model introduced by Davidson and Beaver [1977]. We prove results that give optimal designs, both for the extensions of the Bradley–Terry model and the extensions of the MNL model, and conduct simulations of these models. Finally, we prove results that give optimal designs for the MNL model when the starting design is an orthogonal array constructed using the Rao–Hamming construction, rather than a complete factorial design.

Chapter 1

Introduction and Preliminary Definitions

This thesis is about how best to design choice experiments for three specific models. In this chapter we provide a compendium of relevant definitions and results.

In a *choice experiment*, we present respondents with a series of N *choice sets*. Each choice set consists of two or more *options*. Let there be $m \geq 2$ options in a choice set. The task for each respondent is to make a choice based only on the information provided about the options presented in that choice set. This choice may be to indicate which option they find most appealing or which option is least appealing, or sometimes to indicate simultaneously the 'best' and the 'worst'.

In this thesis, we focus on *forced choice experiments*, where the respondents are compelled to choose one or more of the options in each choice set. The alternative to this is to provide the respondent with the opportunity not to select any of the options presented in the choice set. We do this by adding a *none of these* option to each choice set.

We present the information about each option in the form of *attributes*, which take one of several *levels*. Attributes describe certain features of an option. Let each option be described by k attributes, which take one of ℓ_1, ℓ_2, ..., ℓ_k levels respectively. The experimenter thinks that these features make a contribution to the decision making process. Each option will then consist of a combination of the attribute levels, with one level specified for each attribute. We call a combination of attribute levels an *item*. This is best illustrated in an example.

■ **EXAMPLE 1.0.1.**

Adapted from Phillips et al. [2002]

This experiment was conducted to examine preferences for HIV test methods. Some of the attributes tested, and the levels the attributes may take, are shown in the Table 1.1. We see that there are $k = 4$ attributes: Location, Price, Sample collection, and Timeliness/accuracy. The levels for the Location attribute, for example, are Public clinic, Doctor's office, and Home, so $\ell_1 = 3$. Table 1.2 shows a typical choice set with $m = 2$ options that could be presented to a respondent. □

Once the choices have been collected from respondents we use the data to estimate a choice model, such as the multinomial logit model (MNL model), formally introduced in Section 1.1.

Attributes	Levels
Location	Public clinic
	Doctor's office
	Home
Price	$0
	$10
	$50
Sample collection	Draw blood
	Swab mouth/oral fluids
	Urine sample
Timeliness/accuracy	Results in 1-2 weeks, almost always accurate
	Immediate results, almost always accurate
	Immediate results, less accurate

Table 1.1: Attributes and levels for the HIV test experiment

Attribute	Option 1	Option 2
Location	Public clinic	Home
Price	$10	$50
Sample Collection	Draw Blood	Urine sample
Timeliness/Accuracy	Results in 1-2 weeks, almost always accurate	Immediate results, less accurate

Imagine that you were about to undergo a HIV test.
Which testing method would you prefer? *(tick one only)*
Option 1 ◯ Option 2 ◯

Table 1.2: A typical choice set for the HIV test experiment

We then use this choice model to determine the relative contribution of each of the attributes to the desirability of the items presented in the experiment.

To estimate the model accurately, the researcher needs to ensure that these experiments do not place an undue burden on the respondents. This need to collect as much information as possible in as few choice sets as possible means that the efficient design of the choice sets to be included in the choice experiment is an important consideration.

The goal of this thesis is to extend the current theory on the efficient design of choice experiments. This extension includes optimal design theory that allows for ties, or for the estimation of position effects, or that have smaller starting designs.

In this chapter, we review the existing results that are relevant to the design and analysis of choice experiments. We begin by introducing the basic models for analysing choice experiments. We continue by introducing a method for constructing choice experiments from the standard designs introduced in the appendix of this chapter. We conclude by reviewing the currently known results that give optimal designs for these choice models. The appendices to this chapter review some useful definitions and other results from discrete mathematics, and more specifically, design theory.

1.1 Models for Choice Experiments

In this section, we look at the models that are commonly used to analyse choice experiments. We begin by looking at paired comparisons experiments, in particular the Bradley–Terry model. We then look at some extensions to the Bradley–Terry model, which will form the motivation for later chapters. We conclude this section by looking at the multinomial logit model, which allows for an arbitrary choice set size.

1.1.1 The Bradley–Terry Model

Paired comparisons modelling was made popular by Thurstone [1927] who conducted an experiment to determine the relative seriousness of a set of 19 offences. Thurstone presented pairs of offences to the respondent, and asked them to select the offence that they considered more serious. For example, the author asked respondents the choose between bootlegging and arson, or between homicide and vagrancy. He presented 171 pairs in total.

The model used in Thurstone [1927] was improved upon by Bradley and Terry [1952] by changing the distribution of the error term in the model. This improved model is called the *Bradley–Terry paired comparison model*, or simply the Bradley–Terry model.

The Bradley–Terry model was first used to analyse information obtained during taste testing experiments. In these experiments, the respondent was presented with two food samples, and was asked to indicate which tasted better. We estimate the relative merit of each item using the preferences collected. We denote the *merit* of an item T_i as π_i, and impose the normalising constraint

$$\prod_{i}^{t} \pi_i = 1.$$

We use these merits to find the probability of selecting a particular item T_i when compared to another item T_j. When item T_i appears with item T_j in a choice set, then the probability that item T_i is selected from the choice set is

$$P(T_i|\{T_i, T_j\}) = \frac{\pi_i}{\pi_i + \pi_j}.$$

Bradley and Terry [1952] derive the likelihood function and maximum likelihood estimates for the entries in $\boldsymbol{\pi}$ when testing particular hypotheses, and consider appropriate test statistics. The *likelihood function* is the joint distribution function for the model parameters given a particular sample. That is, if we let $f(\mathbf{x}|\boldsymbol{\theta})$ denote the joint probability density function (PDF) of the independent sample $\mathbf{X} = (X_1, \ldots, X_n)$, then if we observe $\mathbf{x}$ the likelihood function is given by

$$L(\boldsymbol{\theta}|\mathbf{x}) = f(\mathbf{x}, \boldsymbol{\theta}).$$

We can use this likelihood function to find the estimates for the parameters in $\boldsymbol{\theta}$ that fit the data well. We call these values the *maximum likelihood estimators* (MLEs). The MLEs $\hat{\boldsymbol{\theta}}$ for the parameters $\boldsymbol{\theta}$ are the values of $\boldsymbol{\theta}$ that maximise the likelihood function.

In practice, we maximise the log–likelihood function by differentiating with respect to each parameter, and setting the derivative to 0. The resulting system of equations are called the *normal equations.*

We find the MLEs by solving the normal equations simultaneously. We can also find the variances and covariances of the MLEs. The *Fisher information matrix* $I_\xi(\boldsymbol{\theta})_{ij}$ for a design ξ is the inverse of the variance–covariance matrix, and has elements

$$I_\xi(\boldsymbol{\theta})_{ij} = -\mathcal{E}_x\left(\frac{\partial^2 \ln(f(\boldsymbol{x}, \boldsymbol{\theta}))}{\partial\theta_i\theta_j}\right) = \mathcal{E}_x\left(\left(\frac{\partial \ln(f(\boldsymbol{x}, \boldsymbol{\theta}))}{\partial\theta_i}\right)\left(\frac{\partial \ln(f(\boldsymbol{x}, \boldsymbol{\theta}))}{\partial\theta_j}\right)\right).$$

If we let r_{ijb} be the rank given to item T_i in the b^{th} repetition of the comparison between items T_i and T_j, then the likelihood function is given in the next theorem.

■ **THEOREM 1.1.1.**

(Bradley and Terry [1952]) *The likelihood function for the Bradley–Terry model for s repetitions of N pairs of items is given by*

$$L(\boldsymbol{\pi}|\mathbf{r}) = \frac{\prod_i \pi_i^{2s(N-1)-\sum_{j\neq i}\sum_b r_{ijb}}}{\prod_{i<j}(\pi_i + \pi_j)^{n_{\{i,j\}}}},$$

where $n_{\{i,j\}}$ is the number of times the choice set $\{T_i, T_j\}$ is presented in the choice experiment.

□

Both Zermelo [1929] and Ford [1957] consider the convergence requirements for paired comparisons models. Zermelo [1929] examines the convergence requirements for the Thurstone paired comparisons model, and Ford [1957] considers the convergence requirements for the Bradley–Terry model. The assumption that guarantees the convergence of $\boldsymbol{\pi}$ estimates in the Bradley–Terry Model is "if in every possible partition of the objects into two non–empty subsets, some object in the second set has been preferred at least once to some object in the first set" Ford [1957].

The Bradley–Terry model is further developed by [Bradley, 1954a,b, 1955]. In particular, Bradley [1955] considers some large sample properties of the parameter estimates. In Bradley

[1955], Bradley gives the expectation, variance and covariance of the number of times an item is chosen. Bradley continues by giving asymptotic joint distributions for the maximum likelihood estimates of the parameters in the $\boldsymbol{\pi}$ vector. We let $w_{i|C}$ be a selection indicator which takes the value 1 if item T_i is selected from the choice set $C = \{T_i, T_j\}$, and zero otherwise. We also let $w_i = \sum_{C|T_i \in C} w_{i|C}$.

■ **THEOREM 1.1.2.**
(Bradley [1955]) *The selection indicator $w_{i|\{i,j\}}$ is a binomial random variable with expectation $\pi_i(\pi_i + \pi_j)^{-1}$ and variance $\pi_i \pi_j (\pi_i + \pi_j)^{-2}$. Then*

$$\mathcal{E}_\pi(w_i) = \sum_{C|T_i \in C} \frac{\pi_i}{(\pi_i + \pi_j)},$$
$$Var_\pi(w_i) = \sum_{C|T_i \in C} \frac{\pi_i \pi_j}{(\pi_i + \pi_j)^2},$$

and

$$Cov_\pi(w_i, w_j) = \frac{-\pi_i \pi_j}{(\pi_i + \pi_j)^2}.$$ □

■ **THEOREM 1.1.3.**
(Bradley [1955]) *If $\hat{\pi}_1, \ldots, \hat{\pi}_{t-1}$ are the maximum likelihood estimates of $\pi_1, \ldots, \pi_{t-1}$ then the joint limiting distribution of $\sqrt{n(\hat{\pi}_1 - \pi_1)}, \ldots, \sqrt{n(\hat{\pi}_{t-1} - \pi_{t-1})}$ is a normal distribution with zero mean, subject to some regularity conditions.* □

One of the original limitations of the Bradley–Terry model was "where observations do not come from a single population but from distinct but related populations, related in the sense that some populations reasonably may be assumed to have some parameters in common". This observation is made in Bradley and Gart [1962].

Bradley and Gart [1962] introduced some conditions that ensured that the estimators for the parameters remained consistent, and that the asymptotic distributions were unaltered, despite sampling from associated populations. These conditions are:

- Existence of the first three partial derivatives of the distribution function;
- The first three derivatives of the distribution function must be bounded and convergent almost everywhere; and
- The information matrix for the parameter estimates is positive definite,

where we define associated populations to be a set of distinct, but related, populations.

Another restriction that was originally placed on the Bradley-Terry model was that each pair needed to be replicated the same number of times. This restriction is relaxed in Dykstra [1960], who derives likelihood estimates and sets up appropriate hypothesis tests when the number of times each pair appears in the experiment is allowed to vary.

The Bradley-Terry model has some useful properties. The first is that the model is consistent with a set of choice axioms proposed by Luce [1959]. Luce introduced a set of choice axioms based on a probability measure $P \in [0, 1]$ to describe and model the behaviour of a rational individual.

■ **AXIOM 1.1.4.**

Luce [1959] (Luce's Choice Axiom)

Let T be a finite subset of U (the set of all possible alternatives) such that for every $S \subset T$, P_S is defined.

1. *If $P(x, y) \neq 0, 1$, then for all $x, y \in T$, and all $R \subset S \subset T$, $P_T(R) = P_S(R)P_T(S)$, and*
2. *If $P(x, y) = 0$ for some $x, y \in T$, then for every $S \subset T$, $P_T(S) = P_{T-\{x\}}(S - \{x\})$.*

We demonstrate these ideas with an example.

■ **EXAMPLE 1.1.1.**

Consider the choice of the mode of transport used to travel to work. Suppose that there are four options, car, car pool, bus and train. The first two of these modes are private, and the other two modes are public. Then according to the first part of Luce's choice axiom, the probability of choosing to drive a car is

$$P(\text{car}) = P(\text{private}) \times P(\text{car}|\text{private}).$$

According to the second part of the axiom, if $P(\text{train}) = 0$ then

$$P(\text{public}) = P(\text{public but not train}).$$ □

Luce discussed this model in terms of utility. *Utility* is the perceived benefit that the respondent experiences by choosing a particular option. If an option T_1 has a higher utility than another option T_2, then a rational respondent will choose T_1. We can decompose utility into two components, a deterministic component and a stochastic component. That is, the utility of an item T_i experienced by respondent α can be expressed as

$$U_{\alpha i} = V_{\alpha i} + \epsilon_{\alpha i},$$

where $V_{\alpha i}$ is the deterministic term based on observed attributes, and $\epsilon_{\alpha i}$ is the stochastic term, which captures any attributes that are relevant, but not specified in $V_{\alpha i}$.

Utility is linked to the Bradley–Terry model by letting

$$\pi_i = e^{V_i}.$$

The deterministic component of utility, $V_{\alpha i}$, can be further decomposed into a linear combination of attribute levels, and interactions between attributes. In matrix form this would be

$$\boldsymbol{U}_\alpha = B\boldsymbol{\beta}\boldsymbol{X}_\alpha + \boldsymbol{\epsilon}_\alpha,$$

where $\boldsymbol{X}_\alpha$ is the design matrix for respondent α, and $B\boldsymbol{\beta}$ is a matrix of estimable contrasts, with contrast coefficients in B. The goal of a choice experiment is to produce good estimates for the contrasts in $B\boldsymbol{\beta}$, and thus make predictions about respondent behaviour.

Bradley and El-Helbawy [1976] discuss the estimation of contrasts when fitting a Bradley-Terry model, and derive maximum likelihood estimators for the matrix of contrast effects, $B\boldsymbol{\beta}$. The following definition sets up the notation Bradley and El–Helbawy used to estimate contrasts of the attribute effects, which we will use later.

El-Helbawy and Bradley [1978] show that the information matrix for $\sqrt{sN}\hat{\boldsymbol{\pi}}$ can be expressed as

$$I(\boldsymbol{\pi})_{ij} = \sum_{i_1<i_2} \lambda_{\{i_1,i_2\}} E_\pi \left(\left(\frac{\partial \ln f_{\{i_1,i_2\},\alpha}(\mathbf{w},\boldsymbol{\pi})}{\partial \pi_i} \right) \left(\frac{\partial \ln f_{\{i_1,i_2\},\alpha}(\mathbf{w},\boldsymbol{\pi})}{\partial \pi_j} \right) \right),$$

where the summation is over all distinct choice sets in the design.

We define $\Sigma(\boldsymbol{\pi}_a)$ to be the variance–covariance matrix for $\sqrt{sN}\boldsymbol{\gamma}$, where $\boldsymbol{\gamma} = \ln(\boldsymbol{\pi})$. Then El-Helbawy and Bradley [1978] show that the estimates for $B_h\boldsymbol{\gamma}$ are statistically consistent, and find the asymptotic distributions of these estimates.

■ **THEOREM 1.1.5.**
(El-Helbawy and Bradley [1978]) *Given that*

1. *In every partition of the indices* $1, \ldots, t$ *into two non–empty subsets* S_1 *and* S_2 *there exists* $i \in S_1$ *and* $j \in S_2$ *such that* $\lambda_{\{i,j\}} > 0$ *where we define*
$$\lambda_{\{i,j\}} = \lim_{N\to\infty} \frac{n_{\{i,j\}}}{N},$$
for $i \neq j$, *and* $i, j = 1, \ldots, t$,
2. *each element of the probability vector* $\boldsymbol{\pi}$ *is positive, and*
3. $\begin{pmatrix} \mathbf{1}_t \\ \boldsymbol{B}_a \end{pmatrix} \boldsymbol{\gamma} = \mathbf{0}_{a+1}$,

then $\sqrt{N(\hat{\boldsymbol{\gamma}} - \boldsymbol{\gamma})}$ *has a limiting distribution function that is singular, t-variate normal in the space of* $(t-a-1)$ *dimensions with zero mean vector and variance–covariance matrix*

$$\Sigma_1(\boldsymbol{\pi}) = B_h\Sigma(\boldsymbol{\pi})B_h = (C(\boldsymbol{\pi}))^{-1}. \qquad \square$$

El-Helbawy and Bradley [1978] show that the entries of $\Lambda(\boldsymbol{\pi}) = (\Sigma(\boldsymbol{\pi}))^{-1}$ are given by

$$\Lambda(\boldsymbol{\pi})_{ii} = \pi_i \sum_{C|T_i\in C} \lambda_{\{i,i_2\}} \frac{\pi_i}{(\pi_i + \pi_{i_2})^2}, \text{ and}$$
$$\Lambda(\boldsymbol{\pi})_{ij} = -\lambda_{\{i,j\}} \frac{\pi_i\pi_j}{(\pi_i + \pi_j)^2},$$

where the summation is over all choice sets, and $\lambda_{\{i,j\}}$ was defined in Theorem 1.1.5. Under the null hypothesis of equal merits, these entries become

$$\Lambda(\boldsymbol{\pi}_0)_{ii} = \frac{1}{4} \sum_{C|T_i\in C} \lambda_{\{i,i_2\}}, \text{ and}$$
$$\Lambda(\boldsymbol{\pi}_0)_{ij} = -\frac{1}{4}\lambda_{\{i,j\}}.$$

1.1.2 Extensions of the Bradley–Terry Model

Researchers have made several extensions to the Bradley–Terry model to accommodate different situations which arise in investigating choice behaviour. Amongst these are models that incorporate ties and position effects, and loglinear forms for the Bradley–Terry model. We now introduce these extensions. Chapters 2–5 look at these models in greater detail, as well as some generalisations of these models.

Incorporating Ties

As we have mentioned previously, in most discrete choice experiments we force the respondent to choose one of the options presented, or in some cases we allow them to choose none of them. There are some occasions where this is not ideal, and we would like to allow a respondent to find sets of two or more options equally attractive.

Before we discuss how ties could be incorporated into utility theory, let us consider the meaning of a tied preference. In the absence of tied preferences, if two options were equally preferable we would expect that the respondent would choose each of the options with probability 0.5. In the presence of repeated choice sets, this may be interpreted as inconsistency, rather than random selection, on the part of the respondent. It is important to note that even by adding the option of tied preferences, there is still a random element in the choices made by respondents. That is, a respondent may state a tied preference between two options on one occasion, but choose one of the items when presented with the same pair on a different occasion.

A tied preference does not mean that the respondent is choosing both of the options, nor does it mean that are they choosing neither of the options. While this might be the intention when allowing the respondent to state equal preferences, it might not be reason the respondent made that decision. The respondent might choose to tie because they do not want to accept (or reject) either of the options, or perhaps even because they are unable to evaluate the options. From this point on, we assume that a respondent declares a tie only when they find the options equally preferable.

In terms of utility theory, we can interpret this indifference as a result of the utilities of the two options being too close in value to be distinguishable by the respondent. This idea was initially introduced by Glenn and David [1960], who extended the Thurstone–Mosteller model to incorporate ties. Rao and Kupper [1967] were the first authors to incorporate ties into the Bradley–Terry model.

■ **THEOREM 1.1.6.**
(Rao and Kupper [1967]) *Suppose that there exists a utility threshold η such that if the utility differs by less than η then the respondent will declare a tie, and let $\theta = e^{\eta}$. Then the preference probabilities will be*

$$\begin{aligned} P(T_i|\{T_i,T_j\}) &= \frac{\pi_i}{\pi_i+\theta\pi_j}, \\ P(T_j|\{T_i,T_j\}) &= \frac{\pi_j}{\theta\pi_i+\pi_j}, \text{ and} \\ P(\{T_i,T_j\}|\{T_i,T_j\}) &= \frac{\pi_i\pi_j(\theta^2-1)}{(\pi_i+\theta\pi_j)(\theta\pi_i+\pi_j)}. \end{aligned}$$
□

We call this the *Rao–Kupper ties model*.

One failing of the Rao–Kupper model is that it is not consistent with Luce's choice axiom as shown by Davidson [1970]. That is, the Rao–Kupper ties model does not satisfy the criterion that for $P(T_i|\{T_i,T_j\}) \neq 0$, we require

$$\frac{P(T_i|\{T_i,T_j\})}{P(T_j|\{T_i,T_j\})} = \frac{\pi_i}{\pi_j},$$

as is the case in the Bradley–Terry model. Davidson then derived a modification of the Bradley–Terry model that incorporates ties and is consistent with this criterion. Davidson suggests that the probability of a tie between items should be proportional to the geometric mean of the merits of the items that are found to have similar merit (i.e. $\pi_{\{i,j\}} = \nu\sqrt{\pi_i \pi_j}$). As a consequence, when the utilities of the items are similar, the probability of a respondent stating that they are indifferent between the items will be greater. The *Davidson ties model* is introduced in the next theorem.

■ **THEOREM 1.1.7.**
(Davidson [1970]) *Assuming that ν is independent of items T_i and T_j, then the preference probabilities with ties incorporated into the Bradley–Terry model are*

$$\begin{aligned} P(T_i|\{T_i,T_j\}) &= \frac{\pi_i}{\pi_i+\pi_j+\nu\sqrt{\pi_i\pi_j}}, \\ P(T_j|\{T_i,T_j\}) &= \frac{\pi_j}{\pi_i+\pi_j+\nu\sqrt{\pi_i\pi_j}}, \quad \textit{and} \\ P(\{T_i,T_j\}|\{T_i,T_j\}) &= \frac{\nu\sqrt{\pi_i\pi_j}}{\pi_i+\pi_j+\nu\sqrt{\pi_i\pi_j}}. \end{aligned}$$

□

Davidson suggests that $\nu \geq 0$, or rather $1/\nu$ is a measure of how easily the respondent can discriminate between items. If $\nu = 0$, then the respondent will never state a tied preference, since $P(T_i, T_j|T_i, T_j) = \frac{0}{\pi_i+\pi_j+0}$. In this situation, the Davidson ties model simplifies to the Bradley–Terry model. Furthermore, Davidson assumes that if there are no tied preferences stated by any of the the respondents then the Bradley–Terry model should be used instead. We make the same assumption in this thesis.

As ν becomes infinitely large, the probability that the respondent states a tied preference approaches 1. Suppose that $\pi_i = \pi_j = 1$. If $\nu = 0$, then $P(T_i|T_i, T_j) = P(T_j|T_i, T_j) = 0.5$, and $P(T_i, T_j|T_i, T_j) = 0$. If $\nu = 1$, then $P(T_i|T_i, T_j) = P(T_j|T_i, T_j) = P(T_i, T_j|T_i, T_j) = 1/3$. Finally if $\nu = 4$, then $P(T_i|T_i, T_j) = P(T_j|T_i, T_j) = 1/6$, and $P(T_i, T_j|T_i, T_j) = 2/3$. So we see that as ν increases, the probability that the respondent states a tied preference increases as well.

de Dios Ortúzar et al. [2000] presents a paired comparisons experiment with ties to determine the demand for a cycle–way network in Santiago, Chile. While the authors do not elaborate on how these tied preferences are incorporated into the modelling, the survey instrument that they use does give the option of stating that the items presented are equally preferable.

Grutters et al. [2008] also give respondents the option of a tied preference in their paired comparisons experiment to determine the willingness to pay for hearing aids. These authors used a random effects ordered probit model to model the choice behaviour. Such an approach is reasonable in a paired comparisons experiment, but there is no intuitive generalisation to an arbitrary choice set size.

Incorporating Position Effects

Another extension to the Bradley–Terry model that has been considered in the literature is the incorporation of position effects. A position effect models incorporates the effect that the position the item takes within the choice set has on the probability that the item will be selected. Beaver and Gokhale [1975] were the first authors to consider incorporating position effects, which they assumed were additive, into the choice model.

Davidson and Beaver [1977] argue that it is be more natural to consider a multiplicative order effect for choice sets of size 2. Suppose that there is a new parameter $\psi_{ij} \geq 0$ which, when multiplied by the utility of the treatment presented second, inflates or deflates the merit of an item to reflect the effect of the item being presented in the second position of the choice set. Then the value of ψ_{ij} determines how the position effect alters the merit of an item. If $\psi_{ij} < 1$ then by presenting an item in the second position, it is less likely that the respondent will choose the item than if it were presented in the first position, all other things being equal. Conversely, if $\psi_{ij} > 1$ then by presenting the item in the second position, it is more likely that the respondent would choose the item than if it were in the first position, all other things being equal. If $\psi_{ij} = 1$ then there is no position effect at all.

■ **THEOREM 1.1.8.**
(Davidson and Beaver [1977]) *Suppose that the items T_i and T_j are presented in that order. Then the selection probabilities of the Davidson–Beaver position effects model are*

$$\begin{aligned} P(T_i|(T_i,T_j)) &= \frac{\pi_i}{\pi_i + \psi_{ij}\pi_j}, \quad \textit{and} \\ P(T_j|(T_i,T_j)) &= \frac{\psi_{ij}\pi_j}{\pi_i + \psi_{ij}\pi_j}. \end{aligned}$$

□

We call this the *Davidson–Beaver position effects model.* For simplicity, we will assume that the position effect is independent of the items in the choice set. That is, $\psi_{ij} = \psi$ for all $i \neq j$.

Tharp and Marks [1990] undertook a study to determine whether such position effects were present when comparing three different brands of either beer, cars, or furniture. In these experiments the researchers performed an analysis of variance on the partworth utilities, and found no significant position effect. Chrzan [1994] also tested the effect of position, but on the selection of a mail order fashion accessory clubs. He found that position effects did exist when branded alternatives were used.

van der Waerden et al. [2006] also considered the effect of the position of an item on the perceived merit on the item when comparing modes of transport. Like Chrzan [1994], the alternatives were labelled, in this case the labels were bicycle, public transport and car. The authors found a small but significant position effect, with respondents appearing to focus on the items presented later in the choice set.

Incorporating Ties and Position Effects

Davidson and Beaver [1977] extend the Davidson–Beaver position effects model to incorporate ties. The authors used the methods introduced by Rao and Kupper [1967] and Davidson [1970] to incorporate ties. The next theorem gives the extended Davidson ties model.

■ **THEOREM 1.1.9.**
(Davidson and Beaver [1977]) *Suppose that the items T_i and T_j are presented in that order. Then the selection probabilities of the Davidson–Beaver extension of the Davidson ties model to incorporate position effects are*

$$\begin{aligned} P(T_i|(T_i,T_j)) &= \frac{\pi_i}{\pi_i + \psi_{ij}\pi_j + \nu\sqrt{\pi_i\pi_j}}, \\ P(T_j|(T_i,T_j)) &= \frac{\psi_{ij}\pi_j}{\pi_i + \psi_{ij}\pi_j + \nu\sqrt{\pi_i\pi_j}}, \quad \textit{and} \\ P(\{T_i,T_j\}|(T_i,T_j)) &= \frac{\nu\sqrt{\pi_i\pi_j}}{\pi_i + \psi_{ij}\pi_j + \nu\sqrt{\pi_i\pi_j}}. \end{aligned}$$

□

Each of the models presented in this section can be expressed as a loglinear model. Fienberg and Larntz [1976] show that the expected number of times that one option was selected over another has the same maximum likelihood estimates as those for the loglinear model. The loglinear model for the Bradley–Terry model is given by

$$\begin{aligned}\ln(m_{ij1}) &= V_i + A_{ij}, \text{ and} \\ \ln(m_{ij2}) &= V_j + A_{ij},\end{aligned}$$

where $m_{ij1} = n_{ij}p_{ij1}$ is the expected number of times item T_i will be selected when items T_i and T_j are presented in a pair, and V_i is the deterministic component of the utility of item T_i. A_{ij} is a normalising constant, and is the effect of the choice set as a whole in the final model. The loglinear form of the Davidson Ties model is given in Critchlow and Fligner [1991], and is

$$\begin{aligned}\ln(m_{ij1}) &= V_i + A_{ij}, \\ \ln(m_{ij2}) &= V_j + A_{ij}, \text{ and} \\ \ln(m_{ij3}) &= \ln(\nu) + \frac{1}{2}(V_i + V_j) + A_{ij}.\end{aligned}$$

The loglinear form for the Davidson–Beaver position effects model, as introduced by Fienberg [1979], is

$$\begin{aligned}\ln(m_{ij1}) &= V_i + \ln(\psi_1) + A_{ij}, \\ \ln(m_{ij2}) &= V_j + \ln(\psi_2) + A_{ij}, \text{ and} \\ \ln(m_{ij3}) &= \ln(\nu) + \frac{1}{2}(V_i + V_j) + A_{ij}.\end{aligned}$$

Without ties, the position effects model becomes

$$\begin{aligned}\ln(m_{ij1}) &= V_i + \ln(\psi_1) + A_{ij}, \text{ and} \\ \ln(m_{ij2}) &= V_j + \ln(\psi_2) + A_{ij}.\end{aligned}$$

One restriction on the Bradley–Terry model and on all of the models presented in this section is that the respondent compares only two options at a time. It is sometimes more efficient to compare more than two items at a time. To do this, a new model is required. We discuss this in the next section.

1.1.3 The Multinomial Logit Model

One of the disadvantages of the Bradley-Terry model is that it restricts the experiment to paired comparisons. This means that it is not possible to model the selection of one item from a set of three or more items.

Luce [1959] extends the Bradley-Terry model to accommodate the comparison of more than two items. Suppose that we present m items $C = \{T_{i_1}, \ldots, T_{i_m}\}$ to the respondent. Then we can estimate a merit, π_{i_a}, for each of these items. As in the Bradley–Terry model, we let $V_i = \ln(\pi_i)$ be the deterministic part of the utility function. Then the probability that the item $T_i \in C$ is chosen is

$$P(T_i|C) = \frac{\pi_i}{\sum_{a=1}^{m} \pi_{i_a}}.$$

We call this model the *multinomial logit model* (MNL model). In this form, we notice that the Bradley–Terry model is a special case of the MNL model with $m = 2$. The MNL assumes that the unobserved components of the utilities, $\epsilon_{\alpha j}$, are independently and identically distributed with a Type I extreme value distribution with a mean of 0. The probability density function for this random variable is

$$f(\epsilon_{\alpha j}) = e^{-\epsilon_{\alpha j}} e^{-e^{-\epsilon_{\alpha j}}}, \text{ where } -\infty < \epsilon < \infty.$$

McFadden [1973] shows that this model is consistent with Luce's choice axiom.

Burgess and Street [2003] show that the entries in the information matrix for the estimation of the entries in $\boldsymbol{\gamma} = \ln(\boldsymbol{\pi})$ are given by

$$\begin{aligned} \Lambda(\boldsymbol{\pi})_{ii} &= \pi_i \sum_{C|T_i \in C} \frac{\lambda_C\big((\sum_{a=1}^m \pi_{i_a}) - \pi_i\big)}{(\sum_{a=1}^m \pi_{i_a})}, \text{ and} \\ \Lambda(\boldsymbol{\pi})_{ij} &= -\pi_i \pi_j \sum_{C|T_i, T_j \in C} \frac{\lambda_C}{(\sum_{a=1}^m \pi_{i_a})}, \end{aligned}$$

where $\lambda_C = n_C/N$, and n_C is the number of times the choice set $C = \{T_{i_1}, T_{i_2}, \ldots, T_{i_m}\}$ appears in the experiment. Under the null hypothesis of equal merits, these entries become

$$\begin{aligned} \Lambda(\boldsymbol{\pi}_0)_{ii} &= \frac{m-1}{m^2} \sum_{C|T_i \in C} \lambda_C, \text{ and} \\ \Lambda(\boldsymbol{\pi}_0)_{ij} &= -\frac{1}{m^2} \sum_{C|T_i, T_j \in C} \lambda_C. \end{aligned}$$

The MNL model assumes what is called *Independence from Irrelevant Alternatives* (IIA). This means that the probability of choosing an option T_i over another option T_j does not depend on any of the other options in the choice set. This also means that when an option is added or removed, that the change in selection probability is proportional across all options in the choice set. This is called *proportional substitution*. That is, if we change an attribute of item T_p, and denote the probability before the change with superscript 0, and after the change with superscript 1, we have

$$\frac{P_{ni}^1}{P_{nj}^1} = \frac{P_{ni}^0}{P_{nj}^0}.$$

While the IIA assumption is a very useful assumption in terms of modelling ease, it is occasionally an inappropriate assumption. The next example show a case where IIA might reasonably be violated.

■ **EXAMPLE 1.1.2.**

Recall the AIDS test experiment in Example 1.0.1. Suppose that we present the choice set in Table 1.3 to the respondents, and the options are selected in the proportions given in the last row of the table. Notice that the proportion of respondents who selected Option 2 is four times the proportion who selected Option 1.

Now suppose that we add a new option, identical to Option 1 except that the test is performed by Doctor B. We would expect that the selection proportions for Options 1 and 3 would be equal. From the IIA assumption, the proportion for respondents selecting Option 2 must still be four times the proportion selecting Option 1. Then the choice set and selection proportions become those in Table 1.4.

It is unreasonable to suggest that, by adding another generic doctor with identical attributes to the first, half of the respondents who preferred the home test over a doctors appointment would now prefer a doctors appointment over the home test. □

Attribute	Option 1	Option 2
Location	Doctor A	Home
Price	$100	$10
Sample Collection	Draw Blood	Swab mouth
Timeliness/Accuracy	Results in 1-2 weeks, almost always accurate	Immediate results, less accurate
Selection Proportion	20%	80%

Table 1.3: Choice set for AIDS test experiment with selection proportions

Attribute	Option 1	Option 2	Option 3
Location	Doctor A	Home	Doctor B
Price	$100	$10	$100
Sample Collection	Draw Blood	Swab mouth	Draw Blood
Timeliness /Accuracy	Results in 1-2 weeks, almost always accurate	Immediate results, less accurate	Results in 1-2 weeks, almost always accurate
Selection Proportion	16.7%	66.7%	16.7%

Table 1.4: Choice set for AIDS test experiment with additional option and selection proportions

1.2 Choice Designs from Fractional Factorial Designs

In this section, we look at a method for constructing of stated choice experiments from some standard designs. Appendix 1.B provides a review of these standard designs. We will use the designs constructed using this method to form the set of competing designs later in the thesis.

We begin with a *starting design*. The starting design is an $N \times k$ array, usually a factorial design or an orthogonal array. The N rows in this starting design form the first options in each of the N choice sets.

■ **EXAMPLE 1.2.1.**
Consider the experiment introduced in Example 1.0.1. This experiment has 4 3–level attributes. A potential starting design is the OA$[9, 4, 3, 2]$ in Table 1.5. This means that the first option in the last choice set is (2210), which translates to the item described in Table 1.6. □

In order to obtain the other options in the choice set, Burgess and Street [2005] suggest the use of generators, as reviewed in Appendix 1.A. If we have m options in each choice set, then we need $m-1$ further generators to obtain the remaining options from the starting design since we may as well let $\boldsymbol{g}_1 = (00\ldots0)$. In general, a choice set is generated using a set of m generators

$$G = (\boldsymbol{g}_1 = \mathbf{0}, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_m).$$

The next example illustrates this idea.

■ **EXAMPLE 1.2.2.**
Suppose that we have an experiment with $m = 3$ and $\ell_1 = \ell_2 = \ell_3 = \ell_4 = 3$, and that we add two

0	0	0	0
0	1	1	1
0	2	2	2
1	0	1	2
1	1	2	0
1	2	0	1
2	0	2	1
2	1	0	2
2	2	1	0

Table 1.5: An OA$[9, 4, 3, 2]$

Attribute	**Option 1**
Location	Home
Price	$50
Sample Collection	Swab mouth/oral fluids
Timeliness/Accuracy	Results in 1-2 weeks, almost always accurate

Table 1.6: The option described by the last row of the OA in table 1.5

generators to the starting design, as shown in Table 1.7. Let the first of these be $\boldsymbol{g}_2 = (0121)$. Then the first attribute of each item remains unchanged, the second and fourth attributes of each item increase by 1 modulo 3, and the third attribute increases by 2 modulo 3. This gives the second option in each choice set, as given in the second set of columns in Table 1.7. Similarly, we add $\boldsymbol{g}_3 = (1220)$ to the starting design to obtain the third option of each choice set. Then each row describes a choice set of size 3, which we would present in turn to the respondent. □

We can characterise a set of generators by looking at which options differ in which attributes. For example, options 2 and 3 in Example 1.2.2 differ in the first, second and fourth attributes. We define a *difference* $\boldsymbol{d}$ between two options a and b to be a vector of length k, there the q^{th} entry is 1 if the q^{th} attribute takes different levels in these options, and 0 otherwise. So the difference between options 2 and 3 in Example 1.2.2 is $\boldsymbol{d} = (1101)$. We can combine these differences into a single vector containing all $m(m-1)/2$ pairwise differences. We call this the *difference vector*, which is denoted by

$$\boldsymbol{v} = (\boldsymbol{d}_1, \boldsymbol{d}_2, \ldots, \boldsymbol{d}_{m(m-1)/2}).$$

We now find the difference vector for the set of generators in our example.

■ **EXAMPLE 1.2.3.**
Consider Example 1.2.2. There, we constructed a choice experiment using the set of generators

$$G = (\boldsymbol{g}_1 = (0000), \boldsymbol{g}_2 = (0121), \boldsymbol{g}_3 = (1220)).$$

The first and second options differ in attributes 2, 3, and 4, so we express the difference between this pair of options as $\boldsymbol{d} = (0111)$. Similarly, the difference between options 1 and 3 is $\boldsymbol{d} = (1110)$,

Option 1				**Option 2**				**Option 3**			
0	0	0	0	0	1	2	1	1	2	2	0
0	1	1	1	0	2	0	2	1	0	0	1
0	2	2	2	0	0	1	0	1	1	1	2
1	0	1	2	1	1	0	0	2	2	0	2
1	1	2	0	1	2	1	1	2	0	1	0
1	2	0	1	1	0	2	2	2	1	2	1
2	0	2	1	2	1	1	2	0	2	1	1
2	1	0	2	2	2	2	0	0	0	2	2
2	2	1	0	2	0	0	1	0	1	0	0

Table 1.7: The 3^4 choice experiment from Example 1.2.2

and the difference between options 2 and 3 is $\boldsymbol{d} = (1101)$. Thus the difference vector for this set of generators is

$$\boldsymbol{v} = ((0111), (1110), (1101)). \qquad \square$$

■ **EXAMPLE 1.2.4.**

Consider a smaller version of the design in Example 1.2.2 with only two attributes which take three levels each. Then we can use the 3^2 full factorial design as the starting design. Suppose that we use the set of generators

$$G = (\boldsymbol{g}_1 = (00), \boldsymbol{g}_2 = (01), \boldsymbol{g}_3 = (21)).$$

Then we obtain the choice sets in Table 1.8. This set of generators has the difference vector

$$\boldsymbol{v}_1 = ((01), (10), (11)),$$

where the differences are written in lexicographic order. This is acceptable since the order of options within a choice set is immaterial at this point. Table 1.9 gives all of the other distinct sets of generators that have this difference vector and contain 00. Table 1.10 shows all distinct difference vectors that give rise to choice sets without repeated items when $k = 2$ and $\ell_1 = \ell_2 = 3$. A sample set of generators for each difference vector is also given. □

When considering a choice experiment with m options in each choice set, and each option described by k attributes, there are several possible difference vectors. We denote these difference vectors by $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_j\}$. For the set of competing designs, if we assume that there are no repeated items in any single choice set, there are no repeated choice sets in an experiment, and that all distinct choice sets characterised by a particular difference vector appear equally often in the experiment, then we can define a series of constants that describe the choice experiment, as did Burgess and Street [2005]. Let

- $i_{\boldsymbol{v}_j}$ be an indicator of whether all choice sets with difference vector $\boldsymbol{v}_j$ appear in the experiment,
- $c_{\boldsymbol{v}_j}$ be the number of choice sets containing the item $00\ldots0$ with the difference vector $\boldsymbol{v}_j$,
- $x_{\boldsymbol{v}_j;\boldsymbol{d}}$ be the number of times the difference $\boldsymbol{d}$ appears in the difference vector $\boldsymbol{v}_j$,

Option 1		Option 2		Option 3	
0	0	0	1	2	1
0	1	0	2	2	2
0	2	0	0	2	0
1	0	1	1	0	1
1	1	1	2	0	2
1	2	1	0	0	0
2	0	2	1	1	1
2	1	2	2	1	2
2	2	2	0	1	0

Table 1.8: The 3^2 choice experiment from Example 1.2.4

00	01	10
00	01	20
00	02	10
00	02	20
00	01	12
00	02	11
00	01	22
00	02	21

Table 1.9: All sets of generators with difference vector $\boldsymbol{v} = ((01), (10), (11))$

Difference Vector			Sample Generator		
01	01	01	00	01	02
01	10	11	00	01	10
01	11	11	00	01	12
10	10	10	00	10	20
10	11	11	00	10	22
11	11	11	00	11	22

Table 1.10: Difference vectors for an experiment with $k = 2$ and $\ell_1 = \ell_2 = 3$

and let $y_{\boldsymbol{d}}$ be defined as

$$y_{\boldsymbol{d}} = \frac{2}{Nm\prod_{q=1}^{k}(l_q-1)^{i_q}} \sum_j c_{\boldsymbol{v}_j} i_{\boldsymbol{v}_j} x_{\boldsymbol{v}_j;\boldsymbol{d}}.$$

We now see how these constants relate to our example.

■ **EXAMPLE 1.2.5.**
Consider the difference vector $\boldsymbol{v}_1$ from Example 1.2.4. Since there are 8 sets of generators characterised by difference vector $\boldsymbol{v}_1$ with 00 in the first position of the choice set, one for each distinct sets of generators in Table 1.9, we have $c_{\boldsymbol{v}_1} = 8$. If all of the choice sets that are characterised by this difference vector are included in the experiment, then $i_{\boldsymbol{v}_1} = 1$, and if none of the choice sets characterised by $\boldsymbol{v}_1$ are included in the experiment, then $i_{\boldsymbol{v}_1} = 0$. Finally, we have $x_{\boldsymbol{v}_1,(00)} = 0$, $x_{\boldsymbol{v}_1,(01)} = 1$, $x_{\boldsymbol{v}_1,(10)} = 1$, and $x_{\boldsymbol{v}_1,(11)} = 1$. Note that in our class of competing designs either all choice sets characterised by a particular difference vector are included, or none of them are. □

1.2.1 Contrasts

Next, we consider the construction of contrast matrices for choice designs when main effects are of interest. These constructions were introduced in Burgess and Street [2005], and are discussed in more detail in Street and Burgess [2007]. We begin with a matrix of orthogonal polynomial contrast coefficients for an ℓ_q level attribute, denoted by B_{ℓ_q}. For example, a 2 level attribute will have a contrast matrix

$$B_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 & 1 \end{bmatrix},$$

and a 3 level attribute will have a contrast matrix

$$B_3 = \begin{bmatrix} \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix}.$$

These form the building blocks for the contrast matrices when factorial starting designs are used. We now consider the construction of contrast matrices for full factorial starting designs, or subsets of attributes that form a full factorial design. In Chapter 6, we continue this discussion, considering the construction of contrast matrices for fractional factorial starting designs. We will begin with an example, and then look at the general case.

■ **EXAMPLE 1.2.6.**
In this example, we will construct the contrast matrix for the estimation of main effects for the first option in Table 1.8. The first option in this table gives the possible items in lexicographic order. The items form the column labels of the contrast matrix.

Consider the first attribute. As this attribute can take levels 0, 1, or 2, $(0, 1, 2)$ is a row vector that contains the possible levels of the first attribute. Notice that each entry of this levels vector appears three times in a row in the first column of option 1. We can turn this into a vector of levels by taking the Kronecker product of $(0, 1, 2)$ and $\boldsymbol{j}_3^T$, a row vector containing three 1s, to give

$$(0, 1, 2) \otimes \boldsymbol{j}_3^T = (0, 0, 0, 1, 1, 1, 2, 2, 2).$$

This gives the levels that the first attribute takes in each of the first options of the choice experiment. Since the contrast matrix B_3 has columns containing the contrast coefficients corresponding to levels 0, 1, and 2 respectively, the rows of the contrast matrix corresponding to the main effect of the first attribute are given by

$$B_3 \otimes \frac{1}{\sqrt{3}}\boldsymbol{j}_3^T = \begin{bmatrix} \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} \otimes \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$$

$$= \frac{1}{\sqrt{3}} \begin{array}{c} \begin{array}{ccccccccc} 00 & 01 & 02 & 10 & 11 & 12 & 20 & 21 & 22 \end{array} \\ \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} \end{array},$$

where we multiply $\boldsymbol{j}_3^T$ by $\frac{1}{\sqrt{3}}$ to satisfy the scale requirement for contrast coefficients.

Now consider the second attribute. The second attribute also takes levels 0, 1, or 2. The sequence $(0, 1, 2)$ appears three times when reading down the second column of the first option. Again, we use Kronecker products to obtain a row vector corresponding to the levels of the second attribute. We have

$$\boldsymbol{j}_3^T \otimes (0, 1, 2) = (0, 1, 2, 0, 1, 2, 0, 1, 2).$$

Then we can make the same transformation to B_3 to obtain the contrast coefficients corresponding to the main effect of the second attribute. After scaling, this gives

$$\frac{1}{\sqrt{3}}\boldsymbol{j}_3^T \otimes B_3 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \otimes \begin{bmatrix} \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

$$= \frac{1}{\sqrt{3}} \begin{array}{c} \begin{array}{ccccccccc} 00 & 01 & 02 & 10 & 11 & 12 & 20 & 21 & 22 \end{array} \\ \begin{bmatrix} \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} \end{array}.$$

Then we see that the scaled contrast matrix for the estimation of main effects is

$$B_F = \frac{1}{\sqrt{3}} \begin{array}{c} \begin{array}{ccccccccc} 00 & 01 & 02 & 10 & 11 & 12 & 20 & 21 & 22 \end{array} \\ \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} \end{array}$$

$$= \begin{bmatrix} B_3 \otimes \frac{1}{\sqrt{3}}\boldsymbol{j}_3^T \\ \frac{1}{\sqrt{3}}\boldsymbol{j}_3^T \otimes B_3 \end{bmatrix}.$$

□

In general, we consider a complete factorial design F, with k attributes, which take $\ell_1, \ell_2, \ldots, \ell_k$ levels respectively. We can order the combinations of attribute levels in lexicographic order, giving

$$\begin{array}{ccccc} 0 & 0 & \ldots & 0 & 0 \\ 0 & 0 & \ldots & 0 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & \ell_k - 1 \\ 0 & 0 & \ldots & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ \ell_1 - 1 & \ell_2 - 1 & \ldots & \ell_{k-1} - 1 & \ell_k - 1. \end{array}$$

If we order the columns of the contrast matrix to correspond to the items in this order, we obtain a neat structure for the contrast matrix. We use Kronecker products to describe this structure.

Consider the first attribute. The first $\ell_2 \times \ell_3 \times \ldots \times \ell_k$ columns of the contrast matrix will correspond to attribute level 0. The next $\ell_2 \times \ell_3 \times \ldots \times \ell_k$ columns of the contrast matrix will correspond to level 1. This pattern continues until the final $\ell_2 \times \ell_3 \times \ldots \times \ell_k$ columns of the contrast matrix, which correspond to level $\ell_q - 1$. then we can express the rows of the contrast matrix corresponding to the main effect contrasts for the first attribute as

$$B_{\ell_1} \otimes \frac{1}{\sqrt{\ell_2}} \boldsymbol{j}_{\ell_2}^T \otimes \ldots \otimes \frac{1}{\sqrt{\ell_k}} \boldsymbol{j}_{\ell_k}^T,$$

where $\boldsymbol{j}_{\ell_q}$ is a vector containing ℓ_q 1s.

Now consider the second attribute. When written in lexicographic order, we see $\ell_3 \times \ell_4 \times \ldots \times \ell_k$ 0s, followed by $\ell_3 \times \ell_4 \times \ldots \times \ell_k$ 1s, and so on. this sequence is repeated ℓ_1 times. As a result, we may express the rows of the contrast matrix corresponding to the main effect contrasts for the second attribute as

$$\frac{1}{\sqrt{\ell_1}} \boldsymbol{j}_{\ell_1}^T \otimes B_{\ell_2} \otimes \frac{1}{\sqrt{\ell_3}} \boldsymbol{j}_{\ell_3}^T \otimes \ldots \otimes \frac{1}{\sqrt{\ell_k}} \boldsymbol{j}_{\ell_k}^T.$$

We can continue this process until we reach the final attribute which, when written in lexicographic order, appear as $\ell_1 \times \ell_2 \times \ldots \times \ell_{k-1}$ repetitions of $(0, 1, 2, \ldots, \ell_k - 1)^T$. Then we can express the rows of the contrast matrix corresponding to the main effect contrasts for the final attribute as

$$\frac{1}{\sqrt{\ell_1}} \boldsymbol{j}_{\ell_1}^T \otimes \ldots \otimes \frac{1}{\sqrt{\ell_{k-1}}} \boldsymbol{j}_{\ell_{k-1}}^T \otimes B_{\ell_k}.$$

Finally we obtain the main effect contrast matrix

$$B = \begin{bmatrix} B_{\ell_1} & \otimes & \frac{1}{\sqrt{\ell_2}} \boldsymbol{j}_{\ell_2}^T & \otimes & \frac{1}{\sqrt{\ell_3}} \boldsymbol{j}_{\ell_3}^T & \otimes & \ldots & \otimes & \frac{1}{\sqrt{\ell_{k-1}}} \boldsymbol{j}_{\ell_{k-1}}^T & \otimes & \frac{1}{\sqrt{\ell_k}} \boldsymbol{j}_{\ell_k}^T \\ \frac{1}{\sqrt{\ell_1}} \boldsymbol{j}_{\ell_1}^T & \otimes & B_{\ell_2} & \otimes & \frac{1}{\sqrt{\ell_3}} \boldsymbol{j}_{\ell_3}^T & \otimes & \ldots & \otimes & \frac{1}{\sqrt{\ell_{k-1}}} \boldsymbol{j}_{\ell_{k-1}}^T & \otimes & \frac{1}{\sqrt{\ell_k}} \boldsymbol{j}_{\ell_k}^T \\ & & & & & & \vdots & & & & \\ \frac{1}{\sqrt{\ell_1}} \boldsymbol{j}_{\ell_1}^T & \otimes & \frac{1}{\sqrt{\ell_2}} \boldsymbol{j}_{\ell_2}^T & \otimes & \frac{1}{\sqrt{\ell_3}} \boldsymbol{j}_{\ell_3}^T & \otimes & \ldots & \otimes & \frac{1}{\sqrt{\ell_{k-1}}} \boldsymbol{j}_{\ell_{k-1}}^T & \otimes & B_{\ell_k} \end{bmatrix}.$$

Now suppose that we would like to reorder the columns of B to reflect the addition of a generator $\boldsymbol{g}_i$ to the design F. We find this technique is useful in Chapter 6 when we consider the optimal selection of a set of generators when we use a fractional factorial starting design.

The addition of a generator $\boldsymbol{g}_i = (g_{i,1}, g_{i,2}, \ldots, g_{i,k})$ to a design F reorders the levels of each attribute. The next example illustrates this idea.

■ **EXAMPLE 1.2.7.**

Suppose that we would like the order of the columns of B_F to be determined by the items in the second position of the choice sets in the experiment in Table 1.8. We obtained the items in the second position of the choice set by adding a generator to the items presented in the first position. For the first attribute, we left the attribute levels unchanged, so there is no need to change the contrast coefficients in the first two rows of B_F. For the second attribute, we added 1 modulo 3 to the level of that attribute. That is, the 0s became 1s, the 1s became 2s, and the 2s became 0s. We can use a permutation matrix, as introduced in Section 1.A to permute the entries in the levels vector accordingly. Let

$$P_1 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

then

$$\begin{aligned}(0,1,2) \cdot P_1 &= (0,1,2) \cdot \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \\ &= (1,2,0).\end{aligned}$$

We then use Kronecker products to obtain

$$\boldsymbol{j}_3^T \otimes ((0,1,2) \cdot P_1) = (1,2,0,1,2,0,1,2,0),$$

the second column of the second option.

If we post–multiply B_3 by P_1 then we permute the columns of B_3 to reflect the addition of 1 to the column labels. That is,

$$\begin{aligned}B_3P_1 &= \begin{bmatrix} \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix}.\end{aligned}$$

So the contrast coefficients for the main effect of the second attribute are given by

$$\begin{aligned}\frac{1}{\sqrt{3}}\boldsymbol{j}_3^T \otimes (B_3P_1) &= \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \otimes \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} \\ &= \frac{1}{\sqrt{3}} \begin{bmatrix} \scriptstyle 01 & \scriptstyle 02 & \scriptstyle 00 & \scriptstyle 11 & \scriptstyle 12 & \scriptstyle 10 & \scriptstyle 21 & \scriptstyle 22 & \scriptstyle 20 \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix}.\end{aligned}$$

For consistency, when permuting the columns of B_3 for the main effect of the first attribute, we can post–multiply B_3 by $P_0 = I_3$, which corresponds to the addition of 0 in the generator. This gives

$$B_{F+\boldsymbol{g}_2} = \begin{bmatrix} (B_3 P_{g_{2,1}}) \otimes \frac{1}{\sqrt{3}}\boldsymbol{j}_3^T \\ \\ \frac{1}{\sqrt{3}}\boldsymbol{j}_3^T \otimes (B_3 P_{g_{2,2}}) \end{bmatrix}$$

$$= \begin{array}{c} \begin{matrix} 01 & 02 & 00 & 11 & 12 & 10 & 21 & 22 & 20 \end{matrix} \\ \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} \end{array},$$

where $g_{2,1} = 0$, and $g_{2,2} = 1$. □

We can use permutation matrices to capture this reordering. We define the permutation matrix $P_{g_{i,q}}$ by

$$(P_{g_{i,q}})_{x,y} = \begin{cases} 1, & \text{if } x + g_{i,q} = y \mod \ell_q, \\ 0, & \text{otherwise.} \end{cases}$$

This permutation matrix will reorder the columns of B_ℓ to reflect the addition of $g_{i,q}$ to the column labels. This addition need not be modulo ℓ_q, but could be within any finite group that is closed under addition. In Chapter 6, we will also use permutation matrices that reflect the action of addition within GF[ℓ].

This idea can be used to give the contrast matrix for $F + \boldsymbol{g}_i$, $B_{F+\boldsymbol{g}_i}$. We define $B_{F+\boldsymbol{g}_i}$ such that if the contrast coefficients for item T_j are in the j^{th} column of B_F, then the contrast coefficients for the item $T_j + \boldsymbol{g}_i$ will be in the j^{th} column of $B_{F+\boldsymbol{g}_i}$. Then $B_{F+\boldsymbol{g}_i}$ is given by

$$B_{F+\boldsymbol{g}_i} = \begin{bmatrix} B_{\ell_1}P_{g_{i,1}} & \otimes & \frac{1}{\sqrt{\ell_2}}\boldsymbol{j}_{\ell_2}^T & \otimes & \frac{1}{\sqrt{\ell_3}}\boldsymbol{j}_{\ell_3}^T & \otimes & \cdots & \otimes & \frac{1}{\sqrt{\ell_{k-1}}}\boldsymbol{j}_{\ell_{k-1}}^T & \otimes & \frac{1}{\sqrt{\ell_k}}\boldsymbol{j}_{\ell_k}^T \\ \frac{1}{\sqrt{\ell_1}}\boldsymbol{j}_{\ell_1}^T & \otimes & B_{\ell_2}P_{g_{i,2}} & \otimes & \frac{1}{\sqrt{\ell_3}}\boldsymbol{j}_{\ell_3}^T & \otimes & \cdots & \otimes & \frac{1}{\sqrt{\ell_{k-1}}}\boldsymbol{j}_{\ell_{k-1}}^T & \otimes & \frac{1}{\sqrt{\ell_k}}\boldsymbol{j}_{\ell_k}^T \\ & & \vdots & & & & & & & & \\ \frac{1}{\sqrt{\ell_1}}\boldsymbol{j}_{\ell_1}^T & \otimes & \frac{1}{\sqrt{\ell_2}}\boldsymbol{j}_{\ell_2}^T & \otimes & \frac{1}{\sqrt{\ell_3}}\boldsymbol{j}_{\ell_3}^T & \otimes & \cdots & \otimes & \frac{1}{\sqrt{\ell_{k-1}}}\boldsymbol{j}_{\ell_{k-1}}^T & \otimes & B_{\ell_k}P_{g_{i,k}} \end{bmatrix}.$$

1.3 Optimal Designs for Choice Experiments

We now turn our attention to the optimal design for the choice models presented in the previous section. We begin this section by looking at a few criteria for design optimality. We then review some of the results in the literature about the optimal design of choice experiments when the Bradley–Terry model is used, and when the MNL model is used.

Up to this point we have been using $I(\boldsymbol{\pi})$ and $\Lambda(\boldsymbol{\pi})$ to denote the information matrices for the estimation of the entries in $\boldsymbol{\pi}$ and $\boldsymbol{\gamma}$ respectively. We now let $C(\boldsymbol{\pi})$ denote the information matrix for the estimation of those contrasts of the entries in $\boldsymbol{\gamma}$ that are of interest.

1.3.1 Optimal Design Theory

There are many criteria that may be used to define optimal designed experiments. Atkinson et al. [2007] provide a comprehensive list of optimality criteria. In this section, we will consider three of these, D–optimality, E–optimality, and A–optimality. In the remainder of the thesis, we will concentrate on the D–optimality criterion.

All three optimality criteria considered in this section depend on different properties of the variance–covariance matrix. A design is *D–optimal* if it is the design that minimises the generalised variance of the parameter estimates over the set of competing designs $\mathfrak{X}$. That is, it is the design that minimises the determinant of the variance–covariance matrix. This is equivalent to maximising the determinant of the information matrix. A design is *E-optimal* if it is the design that minimises the largest eigenvalue of the variance–covariance matrix. This is equivalent to maximising the smallest eigenvalue of the information matrix. A design is *A–optimal* if it is the design which minimises the trace of the variance–covariance matrix, that is, minimising the sum of the variances.

We can also define *efficiencies* of a design ξ using these criteria. The *D–efficiency* of a design ξ will be

$$D_{\text{eff}}(\xi) = \left(\frac{\det(C_{\text{opt}}^{-1})}{\det(C_{\xi}^{-1})} \right)^{\frac{1}{p}},$$

when there are p parameters to estimate, and we define $\det(C_{\text{opt}}) = \max_{\xi \in \mathfrak{X}} \det(C_{\xi})$. The *$A$–efficiency* of a design ξ will be

$$A_{\text{eff}}(\xi) = \frac{\text{tr}(C_{\text{opt}}^{-1})}{\text{tr}(C_{\xi}^{-1})},$$

when there are p parameters to estimate, and we define $\text{tr}(C_{\text{opt}}^{-1}) = \min_{\xi \in \mathfrak{X}} \text{tr}(C_{\xi}^{-1})$.

These criteria will not always lead to the same choice of design, as the next example shows.

■ **EXAMPLE 1.3.1.**

Suppose that there are three possible designs in a particular class of competing designs $\mathfrak{X}$. These designs are labelled ξ_1, ξ_2, and ξ_3 and have the following variance–covariance matrices.

$$\Sigma_{\xi_1} = \frac{1}{2} \begin{pmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{pmatrix}, \quad \Sigma_{\xi_2} = \frac{1}{13} \begin{pmatrix} 60 & 10 & 50 \\ 10 & 50 & 10 \\ 50 & 10 & 50 \end{pmatrix}, \text{ and } \quad \Sigma_{\xi_3} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{pmatrix}.$$

Then by taking the inverse of each of these variance–covariance matrices, the information matrix for the three parameters to be estimated can be found. We get

$$C_{\xi_1} = \frac{1}{7} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad C_{\xi_2} = \frac{1}{240} \begin{pmatrix} 312 & 0 & -312 \\ 0 & 65 & -13 \\ -312 & -13 & 377 \end{pmatrix}, \text{ and } \quad C_{\xi_3} = \frac{1}{6} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The determinants, and largest eigenvalues of each of these information matrices, as well as the traces of the variance–covariance matrices are given in Table 1.11.

So design ξ_3 maximises $\det(C_{\xi_i})$, design ξ_3 minimises $\text{tr}(C_{\xi_i}^{-1})$, and design ξ_1 minimises the largest eigenvalue of $C_{\xi_i}^{-1}$. Since these designs were the only possible designs in $\mathfrak{X}$, design ξ_1 is the E–optimal design, and ξ_3 is both the D–optimal and A–optimal design. We can also give the relative efficiencies between designs.

Design	$\det(\mathbf{C}_{\xi_i})$	$\mathrm{tr}(\mathbf{C}_{\xi_i}^{-1})$	Largest Eigenvalue of $\mathbf{C}_{\xi_i}^{-1}$
ξ_1	0.02332	10.5	3.5
ξ_2	0.0115	12.31	8.358
ξ_3	0.0416	10.0	6.0

Table 1.11: The D–, A–, and E– values for the designs in Example 1.3.1

$$\begin{aligned} D_{\text{eff}}(\xi_1,\xi_3) &= \left(\frac{0.02332}{0.0416}\right)^{\frac{1}{3}} \\ &= 82.5\% \\ D_{\text{eff}}(\xi_2,\xi_3) &= \left(\frac{0.0115}{0.0416}\right)^{\frac{1}{3}} \\ &= 65.1\% \\ A_{\text{eff}}(\xi_1,\xi_3) &= \frac{10}{10.5} \\ &= 65.0\% \\ A_{\text{eff}}(\xi_2,\xi_3) &= \frac{10}{12.31} \\ &= 81.2\% \end{aligned}$$

□

We see that if one design is optimal based on one criterion, it does not necessarily mean that it will be optimal based on any other criteria. This is especially so in the presence of correlation between parameter estimates, as each criterion treats correlation differently. D–optimality incorporates correlation between the parameter estimates through the use of determinants. E–optimality incorporates covariance by finding the largest eigenvalue of the variance–covariance matrix, which is dependent on the covariances. A–optimality does not incorporate covariances into the criterion at all.

While the selection of a set of contrasts to describe a main effect is a valid consideration, we are able to separate this decision from the choice of design. The next example shows that a change in the contrasts leaves the determinant of the information matrix C unchanged, so long as both sets of contrasts form a basis for the main effects contrast subspace for the corresponding attribute.

■ **EXAMPLE 1.3.2.**

Suppose that we have two sets of orthogonal contrasts, both forming a basis for the main effects contrast subspace for a 4–level attribute. The first set will be the set of orthogonal polynomial contrasts

$$B_{4(1)} = \begin{pmatrix} -\frac{3}{2\sqrt{5}} & -\frac{1}{2\sqrt{5}} & \frac{1}{2\sqrt{5}} & \frac{3}{2\sqrt{5}} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2\sqrt{5}} & \frac{3}{2\sqrt{5}} & -\frac{3}{2\sqrt{5}} & \frac{1}{2\sqrt{5}} \end{pmatrix},$$

with the rows labelled $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, and $\boldsymbol{a}_3$. The second set will be a set of contrasts that treat the levels as a 2^2 design, with the third row being $-\boldsymbol{b}_1\boldsymbol{b}_2$ component–wise. This gives

$$B_{4(2)} = \frac{1}{2}\begin{pmatrix} -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{pmatrix},$$

with the rows labelled $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, and $\boldsymbol{b}_3$.

We can find a linear transformation that maps the rows of $B_{4(1)}$ to the rows of $B_{4(2)}$, and another linear transformation that is the inverse of the first. To do this, we set up the system of equations

$$\begin{aligned} \boldsymbol{a}_1 &= h_{11}\boldsymbol{b}_1 + h_{12}\boldsymbol{b}_2 + h_{13}\boldsymbol{b}_3, \\ \boldsymbol{a}_2 &= h_{21}\boldsymbol{b}_1 + h_{22}\boldsymbol{b}_2 + h_{23}\boldsymbol{b}_3, \\ \boldsymbol{a}_3 &= h_{31}\boldsymbol{b}_1 + h_{32}\boldsymbol{b}_2 + h_{33}\boldsymbol{b}_3. \end{aligned}$$

There are 9 parameters to find. We can use the orthogonality properties of the rows of each of the matrices to find each. To find h_{11} we multiply both sides of the first equation by $\boldsymbol{b}_1^T$

$$\boldsymbol{a}_1.\boldsymbol{b}_1^T = h_{11}\boldsymbol{b}_1.\boldsymbol{b}_1^T + h_{12}\boldsymbol{b}_2.\boldsymbol{b}_1^T + h_{13}\boldsymbol{b}_3.\boldsymbol{b}_1^T.$$

We notice that since the rows of $B_{4(2)}$ are orthogonal, $\boldsymbol{a}_2.\boldsymbol{b}_1^T = \boldsymbol{a}_3.\boldsymbol{b}_1^T = 0$, and as a result of the scaling of the rows, $\boldsymbol{b}_1.\boldsymbol{b}_1^T = 1$. Thus

$$h_{11} = \boldsymbol{a}_1.\boldsymbol{b}_1^T.$$

Through multiplication, we obtain

$$h_{11} = \frac{2}{\sqrt{5}}.$$

We can repeat this process for the other parameters, yielding a system of equations

$$\begin{bmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \boldsymbol{a}_3 \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & -1 \\ \frac{-1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0 \end{bmatrix} \times \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \boldsymbol{b}_3 \end{bmatrix}$$

$$B_{4(1)} = RB_{4(2)}.$$

By matrix multiplication, we can show that R is an orthogonal matrix, that is $R^{-1} = R^T$. Orthogonal matrices have a determinant of 1.

Suppose that the information matrix when using the first contrast matrix is

$$C_1 = B_{4(1)}\Lambda B_{4(1)}^T,$$

and the information matrix when using the second contrast matrix is

$$C_2 = B_{4(2)}\Lambda B_{4(2)}^T.$$

We can use the orthogonal property of R, and the fact that R is square to show that C_1 and C_2 have the same determinant

$$\begin{aligned} \det(C_1) &= \det(B_{4(1)}\Lambda B_{4(1)}^T) \\ &= \det(RB_{4(2)}\Lambda B_{4(2)}^T R^T) \\ &= \det(R)\det(B_{4(2)}\Lambda B_{4(2)}^T)\det(R^T) \\ &= \det(R)\det(C_2)\det(R^T) \\ &= \det(C_2), \end{aligned}$$

as required. □

1.3.2 Optimal Designs for the Bradley–Terry Model

This section reviews the existing theory relating to the optimal design of paired comparison experiments when the Bradley–Terry model is used. Since the Bradley–Terry model is nonlinear in parameters, the optimal design depends on the values of the parameters $\boldsymbol{\beta}$. Two main approaches are used in practice to overcome this problem. The first is to take the view that designs that are optimal for $\boldsymbol{\beta} = \mathbf{0}$ will behave well for other values of $\boldsymbol{\beta}$, and use these designs. Other researchers take a Bayesian approach, and use designs that are optimal over a distribution of values of $\boldsymbol{\beta}$. The prior distributions of the entries in $\boldsymbol{\beta}$ may come from managers expectations, or from pilot testing. We will take the first approach in this thesis.

Quenouille and John [1971] introduce the idea of using a subset of the set of all pairs of items in a choice experiment. The authors establish that by showing a well chosen subset of the pairs of items to the respondents, these pairs can be used to estimate the effects that are of interest. The authors continue by conducting a comparison to see which combination of one or more subsets would be more efficient to estimate certain sets of effects. The authors conclude that for symmetrical designs with binary attributes the set of all of the pairs which differ in all of the attributes could be used to construct designs to estimate main effects.

El-Helbawy and Bradley [1978] build on the information matrix for a factorial paired comparison experiment, derived earlier in the paper, to look at designs for a 2^3 factorial experiment. The authors discuss some optimal designs for the estimation of particular sets of effects. This was achieved by constructing $\Lambda(\boldsymbol{\pi}) = (\Sigma(\boldsymbol{\pi}))^{-1}$ based on which pairs contribute to each effect, and then choosing the set of pairs that minimise the determinant of $\Sigma(\boldsymbol{\pi})$.

Littell and Boyett [1977] discuss the problem of designing an $R \times C$ factorial paired comparison experiment. The class of competing designs contained two designs, one with all pairs of items, and another with only those items that differ in the level of one attribute. The authors find that when testing a single main effect, the design including all pairs where only the attribute of interest differs across the options performs better than a design with all pairs. When testing interaction effects, the authors find that the design with all pairs is more efficient.

El-Helbawy [1984] considers the approaches to the construction of designed experiments that Littell and Boyett [1977], El-Helbawy and Bradley [1978], and Quenouille and John [1971] introduce. The author suggests that some of the approaches work well under some model situations (e.g. 2^k), while other designs work well for the estimation of some sets of effects (e.g. main effects only).

El-Helbawy and Ahmed [1984] consider only 2^k factorial paired comparison experiments. The authors use different classes of competing designs for testing different sets of effects. For example, when testing main effects, two competing designs are used. The first design is the set of all pairs which differ in an even number of attributes, and the second design is the set of all pairs that differ in an odd number of attributes. When testing all of the odd–factor interactions, the author compares the design with all pairs which differ in all of the attributes, and the design which includes pairs with some common attributes.

The authors find that the design consisting of all pairs which differ in all attributes is more efficient in the estimation of main effects, whose contrast coefficients are contained in B_M. The authors also find that the class of designs where the pairs have all but one attribute the same are more efficient in the estimation of higher–order interactions than the design with all pairs of items. The theorem below establishes the former result, where t is the number of distinct items.

■ **THEOREM 1.3.1.**

(El-Helbawy and Ahmed [1984]) *Let ξ be the design for which $\lambda_{\{i,j\}}$ equals $\frac{2}{t}$ or 0 as T_i and T_j differ in the levels of each of the k attributes or not for $i \neq j$ and $i, j = 1, \ldots, t$. Then when the rows of B_M correspond to the $\sum_{j=1}^{k}(l_j - 1)$ main effects we find that ξ will be the A–, D–, and E–optimal design.* □

This is best explained by an example.

■ **EXAMPLE 1.3.3.**

Let $k = 3$. Then the set of all distinct pairs which differ in all attributes is shown in Table 1.12. Each of these pairs will be assigned weight $\lambda = \frac{1}{4}$, and all other pairs will receive a weight of 0 in the design. Then by Theorem 1.3.1, this design is optimal for the estimation of main effects. □

Option 1			**Option 2**		
0	0	0	1	1	1
0	1	1	1	0	0
1	0	1	0	1	0
1	1	0	0	0	1

Table 1.12: Optimal 2^3 design for the estimation of main effects, based on Theorem 1.3.1

Street et al. [2001] show that for a 2^k design, $\Lambda(\boldsymbol{\pi}_0)$ can be expressed as

$$4\Lambda(\boldsymbol{\pi}_0) = \left(\sum_{i=1}^{k} \binom{k}{i} a_{k,i} \right) I_{2^k} - \sum_{i=1}^{k} a_{k,i} D_{k,i},$$

where $D_{k,i}$ is a $(0,1)$ matrix of order 2^k with the rows and columns labelled by the items, with a 1 in position (x, y) if the items x and y have i attributes at different levels. The term $a_{k,i}$ is the weight given to each pair which differ in the levels of i attributes. The authors show that when only main effects are of interest, the determinant of the information matrix is

$$\det(C(\boldsymbol{\pi}_0)_{\mathrm{M}}) = 2^{-k} \left[\sum_{i=1}^{k} a_{k,i} \binom{k-1}{i-1} \right]^k .$$

When main effects plus two–factor interactions are of interest, the determinant is

$$\det(C(\boldsymbol{\pi}_0)_{\mathrm{MT}}) = \left[\frac{1}{2} \sum_{i=1}^{k} a_{k,i} \binom{k-1}{i-1} \right]^k \times \left[\sum_{i=1}^{k} a_{k,i} \binom{k-2}{i-1} \right]^{k(k-1)/2} ,$$

subject to the constraint $2^{k-1} \sum_i \binom{k}{i} a_{k,i} = 1$.

Street et al. [2001] use these expressions to find optimal designs for the estimation of main effects plus two– and three–factor interactions for 2^k paired comparisons experiments. The authors extend the set of competing designs to sets of pairs which differ in the levels of i attributes and prove that the designs which satisfy the condition in Theorem 1.3.1 are still optimal over this new class of competing designs.

■ **THEOREM 1.3.2.**

(Street et al. [2001]) *The D–optimal design for the estimation of main effects and two–factor* *interactions for a 2^k paired comparisons experiment when all other effects are assumed zero is given by:*

$$a_{k,i} = \begin{cases} \left(2^{k-1}\binom{k}{(k+1)/2}\right)^{-1} & i = \frac{k+1}{2} \\ 0 & \textit{otherwise,} \end{cases}$$

if k is odd, and

$$a_{k,i} = \begin{cases} \left(2^{k-1}\binom{k}{k/2}\right)^{-1} & i = \frac{k}{2}, \frac{k}{2}+1 \\ 0 & \textit{otherwise,} \end{cases}$$

if k is even. □

■ **EXAMPLE 1.3.4.**

This example will consider two different values of k, $k = 2$ and $k = 3$. First, let $k = 2$, then the optimal design for the estimation of main effects plus two–factor interactions will consist of all of the distinct pairs which differ in the levels of $\frac{k}{2} = 1$ or $\frac{k}{2} + 1 = 2$ attributes, then $a_{2,0} = 0$, $a_{2,1} = \frac{1}{4}$, and $a_{2,2} = \frac{1}{4}$. This design is shown in Table 1.13.

Now if we let $k = 3$, then the optimal design for the estimation of main effects plus two–factor interactions will consist of all of the pairs which differ in $\frac{k+1}{2} = 2$ attributes, then $a_{3,0} = 0$, $a_{3,1} = 0$, $a_{3,2} = \frac{1}{12}$, and $a_{3,3} = 0$. The optimal design is shown in Table 1.14. □

Option 1	Option 2
0 0	0 1
0 1	1 0
0 1	1 1
1 0	0 0
1 0	1 1
1 1	0 0

Table 1.13: Optimal 2^2 design for the estimation of main effects and two–factor interactions based on Theorem 1.3.2

Optimal designs for 2^k paired comparisons designs are also discussed by Graßhoff and Schwabe [2008]. The authors focus on the optimal designs for experiments with either one or two attributes.

El-Helbawy et al. [1994] consider the optimal design of asymmetric paired comparison experiments when only main effects are of interest, and when all pairwise interactions involving a single factor are of interest. When estimating main effects, the authors compare the design with all distinct pairs of items to the design with all distinct pairs of items that differ in every attribute. When estimating interactions involving a single factor, the authors compare the design with all distinct pairs of items with the design with all distinct pairs of items that differ in only the attribute involved in all of the interactions.

■ **THEOREM 1.3.3.**

(El-Helbawy et al. [1994]) *Consider a $\ell_1 \times \ell_2 \times \ldots \times \ell_k$ factorial paired comparison experiment. Assuming that there are no interactions present, and that B_h consists of the main effects, then the design consisting of all distinct pairs where the options differ in all of the attributes will be A–, D–, and E–optimal in the design space.* □

Option 1			Option 2		
0	0	0	0	1	1
0	0	0	1	0	1
0	0	0	1	1	0
0	0	1	1	0	0
0	1	0	0	0	1
0	1	1	1	0	1
1	0	0	0	1	0
1	0	1	1	1	0
1	1	0	0	1	1
1	1	1	0	0	1
1	1	1	0	1	0
1	1	1	1	0	0

Table 1.14: Optimal 2^3 design for the estimation of main effects and two–factor interactions based on Theorem 1.3.2

We illustrate this theorem with an example.

■ **EXAMPLE 1.3.5.**
Consider a 2×3 factorial experiment. Then the optimal design will be the set of all distinct pairs where the items differ in all attributes. For this experiment, the design in Table 1.15 will be optimal. □

Option 1		Option 2	
0	0	1	2
0	1	1	0
0	2	1	1
1	0	0	2
1	1	0	0
1	2	0	1

Table 1.15: Optimal 2×3 factorial design for for the estimation of main effects based on Theorem 1.3.3

Other research into the optimal design of paired comparisons research focuses on continuous designs. van Berkum [1987] proves results that give continuous optimal designs for the estimation of main effects, main effects plus two–factor interactions, and for all quadratic effects. Graßhoff et al. [2003] prove results that give D–optimal continuous designs in the presence of a so–called profile constraint. That is, the authors consider designs where subsets of $S \leq k$ attributes appear in choice sets. Subsequently Graßhoff et al. [2007] proves results that give some small exact D–optimal designs for the estimation of main effects plus two–factor interactions in 2^k experiments.

1.3.3 Optimal Designs for the MNL Model

We now turn our attention to the optimal design of experiments that use the MNL model. Since the MNL model is nonlinear in parameters, the optimal design depends on the values of the parameters $\boldsymbol{\beta}$. Like the Bradley–Terry model, in this thesis we take the view that designs that are optimal for $\boldsymbol{\beta} = \mathbf{0}$ will behave well for other values of $\boldsymbol{\beta}$, and use these designs.

The optimal design results discussed here will be based on the constructions considered in Section 1.2. The class of competing designs will include those designs that include all distinct choice sets with a particular difference vector, or none of the distinct choice sets with a particular difference vector, as stated in Section 1.2.

There are several results on the optimal design of 2^k choice experiments when the MNL model is used. The following theorem gives a method of finding the optimal design for the estimation of main effects, assuming equal selection probabilities. For a 2^k experiment, we define d_{ij} to be the sum of the entries in the $(i, j)^{\text{th}}$ difference.

■ **THEOREM 1.3.4.**
(Burgess and Street [2003]) *The D–optimal design for testing main effects only, when all other effects are assumed to be zero, is given by choice sets in which, for each $\boldsymbol{v}_j$ present,*

$$\sum_{i=1}^{m-1} \sum_{j=i+1}^{m} d_{ij} = \begin{cases} \frac{(m^2-1)k}{4} & m \text{ odd}, \\ \frac{m^2 k}{2} & m \text{ even}, \end{cases}$$

and there is at least one $\boldsymbol{v}_j$ with a non–zero $a_{\boldsymbol{v}_j}$; that is, the choice experiment is nonempty. □

■ **EXAMPLE 1.3.6.**
Consider a 2^3 experiment the estimation of main effects only, and with $m = 3$ options per choice set. In this situation, the optimal design will have $\sum_{i=1}^{2} \sum_{j=i+1}^{3} d_{ij} = 6$. The difference vectors $\boldsymbol{v}_i$ that satisfy this are $(011, 101, 110)$, $(001, 110, 111)$, $(010, 101, 111)$, and $(100, 011, 111)$. Then the set of distinct choice sets which have some or all of these difference vectors will form an optimal design. This means that there are 15 possible optimal designs. One such design is shown in Table 1.16. □

Option 1	**Option 2**	**Option 3**
0 0 0	0 1 1	1 1 0
0 0 1	0 1 0	1 1 1
0 1 0	0 0 1	1 0 0
0 1 1	0 0 0	1 0 1
1 0 0	1 1 1	0 1 0
1 0 1	1 1 0	0 1 1
1 1 0	1 0 1	0 0 0
1 1 1	1 0 0	0 0 1

Table 1.16: Optimal 2^3 factorial design for the estimation of main effects based on Theorem 1.3.4

If in addition to the main effects, we also want to estimate interactions between pairs of attributes, then the next theorem provides a method of finding optimal designs. We define y_i to be the sum of the $y_{\boldsymbol{d}}$s where the difference $\boldsymbol{d}$ has i non–zero entries; that is, the proportion of choice sets which contains each pair with i attributes different.

■ **THEOREM 1.3.5.**
(Burgess and Street [2003]) *The D–optimal design for testing main effects plus two–factor interactions when all other effects are assumed to be zero, is given by*

$$y_i = \begin{cases} \frac{m(m-1)}{2^k}\binom{k+1}{k/2}^{-1} & m \text{ even, and } i = k/2, k/2+1 \\ \frac{m(m-1)}{2^k}\binom{k}{(k+1)/2}^{-1} & m \text{ odd, and } i = (k+1)/2 \\ 0 & \text{otherwise,} \end{cases}$$

when this results in nonzero y_is that correspond to difference vectors that actually exist. □

■ **EXAMPLE 1.3.7.**
Consider the 2^3 experiment with $m = 3$ presented in Example 1.3.6. In this situation, the optimal design for the estimation of main effects plus two–factor interactions will be given by

$$y_i = \frac{m(m-1)}{2^k}\binom{k}{(k+1)/2}^{-1} = \frac{3 \times 2}{8}\binom{3}{2}^{-1} = 0.25 \qquad \text{if } i = \frac{k+1}{2} = 2.$$

The values of y_i tell us which difference vectors can be used to obtain an optimal design. All pairs of items in each choice set must differ in two of the attributes, since $y_i = 0$ for $i \neq 2$. Each pair that differs in the levels of two attributes appears in $y_i = \frac{1}{4}$ of the choice sets. Then all choice sets with the difference vector $(011, 101, 110)$ will form an optimal design (since each difference has two non–zero elements). This gives the design in Table 1.17. □

Option 1	**Option 2**	**Option 3**
0 0 0	0 1 1	1 0 1
0 0 1	0 1 0	1 0 0
0 1 0	0 0 1	1 1 1
0 1 1	0 0 0	1 1 0
1 0 0	1 1 1	0 0 1
1 0 1	1 1 0	0 0 0
1 1 0	1 0 1	0 1 1
1 1 1	1 0 0	0 1 0

Table 1.17: Optimal 2^3 factorial design for the estimation of main effects and two–factor interactions based on Theorem 1.3.5

Although 2^k designs are quite useful, especially in screening experiments, sometimes we need attributes which have more levels. Results on the optimal asymmetric designs exist when only main effects are of interest. To date, there are no results on the optimal design of asymmetric experiments when main effects plus higher order interactions are of interest. Before we introduce a result giving optimal designs for asymmetric experiments, we look at an example of such an asymmetric design.

■ **EXAMPLE 1.3.8.**
Consider a 2×4 factorial experiment with $m = 4$ options per choice set. Let $\boldsymbol{g}_1 = \mathbf{0}$, $\boldsymbol{g}_2 = (13)$,

$\boldsymbol{g}_3 = (02)$, and $\boldsymbol{g}_4 = (11)$. Then the difference vector for this set of generators is

$$\boldsymbol{v} = (01, 01, 11, 11, 11, 11).$$

The first attribute contains the difference 1 four times, and the difference 0 twice. For the second attribute, $g_{i_1q} - g_{i_2q}$ equals $\pm 1 \mod 4$ four times, and $2 \mod 4$ twice, so each non–zero difference modulo ℓ_q appears four times. This design is given in Table 1.18. □

Burgess and Street [2005] show that, for a choice experiment with k attributes, $\Lambda(\boldsymbol{\pi}_0)$ becomes

$$\Lambda(\boldsymbol{\pi}_0) = \frac{m-1}{m^2} z I_L - \frac{1}{m^2} \sum_{\boldsymbol{d}} \frac{y_{\boldsymbol{d}} D_{\boldsymbol{d}}}{\prod_{j=1}^{k} (\ell_j - 1)^{i_j}},$$

where $z = \sum_j c_{\boldsymbol{v}_j} a_{\boldsymbol{v}_j}$ and $D_{\boldsymbol{d}}$ is a $(0,1)$ matrix of order L with rows and columns labelled by the items, with a 1 in position (x, y) if items x and y have difference $\boldsymbol{d}$, and 0 otherwise. The authors show that the determinant of the information matrix for the estimation of main effects only is

$$\begin{aligned}\det(C(\boldsymbol{\pi}_0)_{\mathrm{M}}) &= \prod_{q=1}^{k} \left[\frac{1}{m^2} \sum_{\boldsymbol{d}} y_{\boldsymbol{d}} \left(1 - \frac{1}{(1-\ell_q)^{i_q}} \right) \right]^{\ell_q - 1} \\ &= \prod_{q=1}^{k} \left[\frac{2\ell_q}{m^3(\ell_q - 1)} \sum_j c_{\boldsymbol{v}_j} a_{\boldsymbol{v}_j} \sum_{\boldsymbol{d}|i_q=1} x_{\boldsymbol{v}_j;\boldsymbol{d}} \right]^{\ell_q - 1},\end{aligned}$$

subject to the constraint $\prod_{q=1}^{k} \frac{\ell_q z}{m} = 1$. Since $\sum_{\boldsymbol{d}|i_q=1} x_{\boldsymbol{v}_j;\boldsymbol{d}}$ is the number of non–zero differences for the q^{th} attribute, we expect that this sum will be maximised when the number of zero differences is as small as possible. Then the optimal designs for the estimation of main effects should consist of choice sets where each pair of options in the choice set differ in as many of the attributes as possible, for a given m, ℓ_1, $\ell_2, \ldots, \ell_k$.

Using the construction method and set of competing designs described in Section 1.2, Burgess and Street [2005] find optimal designs for the estimation of main effects in an asymmetric experiment.

■ **THEOREM 1.3.6.**

(Burgess and Street [2005]) *Let F be the complete factorial for k attributes where the q^{th} attribute has l_q levels. Suppose that a set of m generators $G = \{\boldsymbol{g}_1 = \boldsymbol{0}, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_m\}$ such that $\boldsymbol{g}_i \neq \boldsymbol{g}_j$ for $i \neq j$. Suppose that $\boldsymbol{g}_i = (g_{i1}, g_{i2}, \ldots, g_{ik})$ for $i = 1, \ldots m$ and suppose that the multiset of differences for attribute q $\{\pm(g_{i_1q} - g_{i_2q}) | 1 \leq i_1, i_2 \leq m, i_1 \neq i_2\}$ contains each non–zero difference modulo ℓ_q equally often. Then the choice sets given by the rows of $F + \boldsymbol{g}_1, F + \boldsymbol{g}_2, \ldots, F + \boldsymbol{g}_m$ for one or more sets of generators G, are optimal for the estimation of main effects only, provided that there are as few zero differences as possible in each choice set.* □

■ **EXAMPLE 1.3.9.**

When the MNL model is used, the design considered in Example 1.3.8 is optimal for the estimation of main effects only. □

Burgess and Street [2005] also show that the determinant of the information matrix for the estimation of main effects plus two–factor interactions is given by

$$\begin{aligned}\det(C(\boldsymbol{\pi}_0)_{\mathrm{MT}}) = &\prod_{q=1}^{k} \left[\frac{1}{m^2} \sum_{\boldsymbol{d}} y_{\boldsymbol{d}} \left(1 - \frac{1}{(1-\ell_q)^{i_q}} \right) \right]^{\ell_q - 1} \\ &\times \prod_{q_1=1}^{k-1} \prod_{q_2=q_1+1}^{k} \left[\frac{1}{m^2} \sum_{\boldsymbol{d}} y_{\boldsymbol{d}} \left(1 - \frac{1}{(1-\ell_{q_1})^{i_{q_1}} (1-\ell_{q_2})^{i_{q_2}}} \right) \right]^{(\ell_{q_1}-1)(\ell_{q_2}-1)}.\end{aligned}$$

Option 1		Option 2		Option 3		Option 4	
0	0	1	3	0	2	1	1
0	1	1	0	0	3	1	2
0	2	1	1	0	0	1	3
0	3	1	2	0	1	1	0
1	0	0	3	1	2	0	1
1	1	0	0	1	3	0	2
1	2	0	1	1	0	0	3
1	3	0	2	1	1	0	0

Table 1.18: Optimal 2×4 factorial design for the estimation of main effects based on Theorem 1.3.6

While being able to find optimal designs is desirable, researchers also need to take respondent burden into account. That is, even if a design is D–optimal, a design with too many choice sets will place a large burden on respondents. Therefore small, near–optimal designs are also useful in practice. Graßhoff et al. [2004] provide a discussion of the optimality of designs that are constructed from orthogonal arrays and from Hadamard matrices for $m = 2$.

Street et al. [2005] and Street and Burgess [2007] also investigate different methods of obtaining small near–optimal designs for choice experiments with arbitrary choice set size. The authors found that by using a fractional factorial starting design they consistently obtained efficient designs that allowed for the independent estimation of main effects, or main effects plus two–factor interactions. Street and Burgess [2007] found that the SAS macros introduced by Kuhfeld [2005], and the so–called L^{MA} method introduced by Louviere et al. [2000] also provide near–optimal designs, but in general do not allow for the independent estimation of the effects of interest.

1.4 Thesis Outline

Chapter 2 examines the Davidson ties model in more detail when $m = 2$. We find expressions for the normal equations, and the information matrix when this model is used. We use this information matrix to prove results that allow researchers to find optimal designs for the estimation of main effects and ν where attributes take any combination of levels. We also prove results for finding optimal designs for the estimation of main effects plus two–factor interactions and ν for 2^k experiments. We then conduct simulations to determine the ability of the designs generated from these results to estimate main effects and ν. We also simulate the designs when main effects plus two–factor interactions and ν are of interest.

In Chapter 3, we introduce a generalisation of the Davidson ties model that allows an arbitrary choice set size. This generalisation is analogous to the generalisation of the Bradley–Terry model to obtain the MNL model. In this chapter, we derive the normal equations for the MLEs, and the information matrix for the estimation of the contrasts in $B_h\boldsymbol{\gamma}$. We then prove results that give the optimal designs for the estimation of main effects and ν where attributes take any combination of levels. We also prove results that give the optimal design for the estimation of main effects plus two–factor interactions and ν for 2^k experiments. Again, we use simulations to investigate the ability of the designs generated by these results to estimate main effects and ν, as well as to estimate main effects plus two–factor interactions and ν.

In Chapter 4 we derive the information matrix for the estimation of the contrasts in $B_h\boldsymbol{\gamma}$ and position effects when the Davidson–Beaver position effects is used and $m = 2$. We then prove results that give optimal designs for the estimation of main effects and position effects where attributes may have any number of levels. We also prove results that give optimal designs for the estimation of main effects plus two–factor interactions and position effects for 2^k experiments. Once again, we use simulations to investigate the ability of the designs generated by these results to estimate main effects and position effects, as well as to estimate main effects plus two–factor interactions and position effects.

In Chapter 5 we introduce a generalisation of the Davidson–Beaver position effects model that allows for an arbitrary choice set size. We derive normal equations for the MLEs and derive an expression for the information matrix for the estimation of main effects and position effects. We then use this information matrix to prove a general result for finding optimal designs for the estimation of main effects and position effects. We also find an expression for the information matrix for the estimation of main effects, two–factor interactions and position effects. In addition, we discuss designs that are generated by embedding an orthogonal array into a complete Latin square. Once again, we use simulations to investigate the ability of the designs generated by these results to estimate main effects and position effects, as well as to estimate main effects plus two–factor interactions and position effects.

In Chapter 6, we consider designs generated from fractional factorial designs for an arbitrary choice set size. Specifically, we consider symmetric designs with a prime power number of levels and constructed using the Rao–Hamming method, introduced in Section 1.B.3. The benefit of using a fractional factorial design as a starting design is that we present fewer choice sets to the respondent, while maintaining design efficiency. We derive an expression for the information matrix for the estimation of main effects, which will be used to prove optimality results for the estimation of main effects.

In Chapter 7, we provide a summary of the results proved in this thesis, and discuss future research directions arising from this work.

1.A Basic Algebraic Results

The optimal designs for choice experiments rely on the properties of a number of algebraic structures. In this section, we review some results from set theory, group theory, and linear algebra that we use to construct and describe choice designs, and to find optimal designs.

1.A.1 Set Theory

To describe both the choice sets that are presented to the respondent, and the transformation of one set of items into another, we need to introduce some terminology from set theory. The first notion is that of a set. We define a *set* $\{x_1, \ldots, x_n\}$ to be a collection of elements, which could either be finite or infinite.

In Chapters 4 and 5, we consider experiments where the order of presentation is deemed to be important. Thus we need to describe the choice sets in a way that emphasises the importance of order. An *ordered set* $(x_1, \dots, x_n)$ is a sequence of elements that is distinguished both by the identity of the elements and the order of those elements, that is (a, b) is not identical to (b, a), unless $b = a$. We enclose unordered sets in braces {} and ordered sets in parentheses ().

Another common task in the construction of choice experiments is to make a new design from an existing design by adding some number to each of the levels. We will use generators to describe this transformation. A *generator* $\boldsymbol{g}$ is a $1 \times n$ vector which, when added to an n–set $\{x_1, x_2, \dots, x_n\}$ modulo ℓ forms a new n–tuple, where the i^{th} entry is the sum of the i^{th} entry in $\boldsymbol{g}$ and the i^{th} entry in the original set.

■ **EXAMPLE 1.A.1.**
If we have $\boldsymbol{x} = \{1, 2, 4\}$ and add the generator $\boldsymbol{g} = \{2, 2, 1\}$ modulo 5, then

$$\begin{aligned} \boldsymbol{x} + \boldsymbol{g} &= \{1, 2, 4\} + \{2, 2, 1\} \\ &= \{1 + 2, 2 + 2, 4 + 1\} \\ &= \{3, 4, 0\}. \end{aligned}$$

□

To describe the generators used to make new designs from old, we need to introduce the concept of a multiset, which allows elements to appear more than once in the set. This is necessary because generators do not necessarily have distinct entries. Bogart [2000] define an r–element *multiset* chosen from a set S to be an ordered pair (S, f), where S is a set and f is a function from S to the nonnegative integers such that the sum of the values of $f(x)$ for all x in S is m. The number $f(x)$ is called the *cardinality* of x in S.

■ **EXAMPLE 1.A.2.**
Suppose that we have the set $S = \{0, 1, 2, 3, 4\}$, and the function $f(x)$ is given by

$$\begin{aligned} f(0) &= 2 & f(3) &= 1 \\ f(1) &= 0 & f(4) &= 0 \\ f(2) &= 1, \end{aligned}$$

then the multiset defined by this set and function $f(x)$ is

$$A = \{0, 0, 2, 3\}.$$

Notice that the sum of the cardinalities is $m = 4$, the number of items in the multiset. □

Finally, we need to modify the definition of union to allow for multiple entries in the set. We will call this type of union a *strong union*. The strong union between two multisets $A = (S_1, f_1)$ and $B = (S_2, f_2)$ is the multiset $A\&B = (S_1 \cup S_2, f_1 + f_2)$. That is, the cardinality of an item in the strong union is the sum of the cardinalities of the item in the multisets A and B. We illustrate this with an example.

■ **EXAMPLE 1.A.3.**
Suppose that we have two multisets $A = \{a, a, b, b, b, c\}$ and $B = \{b, c, c\}$. Then the strong union between those multisets is

$$A\&B = \{a, a, b, b, b, b, c, c, c\},$$

where there are $2+0 = 2$ as, $3+1 = 4$ bs, and $1+2 = 3$ cs in the strong union of the multisets. □

We denote the strong union between the sets $A_1, A_2, \dots, A_n$ by $\mathop{\&}\limits_{i=1}^{n} A_i$.

1.A.2 Group Theory

The construction of many of the common types of designed experiments depends on concepts in group theory. Choice experiments are no exception. This section introduces the group theory required to discuss the construction of choice designs.

We begin by introducing the concept of a group. Chouinard et al. [2007] defines a *group* $(P, \times)$ as a set P, and a binary operation $\times$ on P, that satisfies the following conditions

- if $x, y \in P$, then $x \times y \in P$;
- an identity element $e \in P$ exists: $x \times e = e \times x = x$ for all $x \in P$;
- $\times$ is associative: $x \times (y \times z) = (x \times y) \times z$ for all $x, y, z \in P$;
- every element $x \in P$ has an inverse, an element x^{-1} for which $x \times x^{-1} = x^{-1} \times x = e$.

Later, we will consider a group where the set P which contains all of the items that could be presented to a respondent in a choice task. The binary operation allows us to transform one item into another in a systematic way.

On occasion, we may not wish to consider every possible item in P, but would like to retain the useful properties of a group. Then we may consider a *subgroup* S of a group P based on a subset of the items in P. For S to be a subgroup, the elements of a subgroup S need to satisfy the criteria for a group under the original group operation, $\times$ (Macdonald [1968]).

One example of a group is a vector space. A *vector space* $(V, +)$ consists of a set of vectors such that if $\boldsymbol{a}, \boldsymbol{b} \in V$ then $\boldsymbol{a} + \boldsymbol{b} \in V$. The $+$ operation on V satisfies all of the properties given by Chouinard et al. [2007]. In addition, we can define scalar multiplication in V so that if $\boldsymbol{a} \in V$ then $\lambda \boldsymbol{a} \in V$, where λ is a scalar. We are also able to define subspaces for vector spaces.

In a vector space, we are also able to define a set of linearly independent vectors from V (or a subspace of V) such that any vector in V (or a subspace of V) can be expressed as a linear combination of this set of vectors. Such a set of vectors is said to be a *basis* of the vector space V (or a subspace of V).

■ **EXAMPLE 1.A.4.**
Consider the vectors in $\mathbb{R}^3$. Then the vectors

$$(1, 0, 0), (0, 1, 0), \text{ and } (0, 0, 1)$$

form a basis of $\mathbb{R}^3$. That is, any vector in $\mathbb{R}^3$ can be expressed as a linear combination of the above three vectors. If we restrict the third entry in a vector to be equal to 0, then we generate a subspace of $\mathbb{R}^3$. The first two vectors will then form a basis for this subspace. □

1.A.3 Finite Fields

Often it is useful to work with a set P and two operators, such as addition and multiplication. If certain properties are satisfied then such a structure is known as a *field*. Street and Street [1986] define a field $\{P, +, \times\}$ as a set, F, closed under two operations, addition (denoted by $+$) and multiplication (denoted by $\times$). Both of these field operations are associative and commutative, and multiplication distributes over addition. Identity elements exist for both of these operations, denoted by 0 for addition and by 1 for multiplication, where $0 \neq 1$. In a field every element

$a \in F$ has an additive inverse $(-a)$ and every non–zero element $a \in F$ has a multiplicative inverse (a^{-1}).

For a prime power $\ell = s^n$, where s is a prime number, we call this field a *Galois field* of order ℓ, denoted by $GF[\ell]$. Street and Street [1986] define a Galois field of order ℓ as a field that can be represented by the set of residue classes of polynomials over $GF[s]$ modulo $f(x)$. The function $f(x)$ is called the *irreducible polynomial*, and must have no factors of the form $\alpha^m - 1$ for $m < n$. The non–zero elements of $GF[\ell]$ form a cyclic multiplicative group, with $\alpha^{\ell-1} = 1$, where α is a root of $f(x)$. We use $f(x)$ to construct the field. The elements of the Galois field can be denoted by 0, 1, α, α^2, $\ldots$, $\alpha^{\ell-2}$.

To find the irreducible polynomial, we look at all of the quadratic functions (since $n = 2$) modulo $s = 2$, checking each to see if the quadratic function can be factorised. We have

$$\alpha^2 = \alpha \times \alpha, \qquad \alpha^2 + 1 = (\alpha + 1)^2, \qquad \alpha^2 + \alpha = \alpha(\alpha + 1),$$

and $\alpha^2 + \alpha + 1$, which cannot be factorised. Then $f(\alpha) = \alpha^2 + \alpha + 1$ is the only irreducible polynomial for $n = s = 2$. Then we set $f(\alpha) = 0$, and obtain $\alpha^2 = \alpha + 1$, since the coefficients are integers modulo $s = 2$.

If $GF[\ell]$ exists, then we can express the field in terms of an n–tuple containing elements of $GF[s]$. We use the irreducible polynomial to construct the n–tuple. Let $(c_0, c_1, \ldots, c_{n-1})$ denote

$$c_0 + c_1\alpha + c_2\alpha^2 + \ldots + c_{n-1}\alpha^{n-1}$$

where $c_0, c_1, \ldots, c_{n-1}$ are integers modulo s. Then we generate n–tuples by substituting $f(\alpha) = 0$ into the expression for each element, and simplify modulo p, to obtain a polynomial of order $n-1$. Then we can represent the levels as $(0, 0, \ldots, 0), (0, 0, \ldots, 0, 1), \ldots, (s-1, s-1, \ldots, s-1)$. The next example demonstrates both of these representations, and field operations within $GF[\ell]$.

■ **EXAMPLE 1.A.5.**

Let $\ell = 4$. Then, since $n = 2$ and $s = 2$, a Galois field of order 4 exists. The elements in this field are 0, 1, α, and α^2, where $\alpha^2 = \alpha + 1$. This Galois field has the addition table and multiplication tables shown Tables 1.19(a) and 1.19(b) respectively.

$+$	0	1	α	α^2
0	0	1	α	α^2
1	1	0	α^2	α
α	α	α^2	0	1
α^2	α^2	α	1	0

$\times$	0	1	α	α^2
0	0	0	0	0
1	0	1	α	α^2
α	0	α	α^2	1
α^2	0	α^2	1	α

Table 1.19: Addition (a) and multiplication (b) tables for $GF[4]$ - multiplicative notation

We now look at an alternative representation of the elements in $GF[4]$, by considering the coefficients of α and 1 in the expression for each element. The element 0 could be represented by $0 \times \alpha + 0 \times 1$, the element 1 could be represented by $0 \times \alpha + 1 \times 1$, the element α could be represented by $1 \times \alpha + 0 \times 1$, and the element $\alpha^2 = \alpha + 1$ could be represented be $1 \times \alpha + 1 \times 1$. Then we obtain the mapping in Table 1.20, where the first entry in the pair is the coefficient of α, and the second entry is the coefficient of 1. This gives the addition and multiplication tables in Tables 1.21(a) and 1.21(b) respectively.

Multiplicative	**Pairs**
0	00
1	01
α	10
$\alpha^2 = \alpha + 1$	11

Table 1.20: Relationship between two representations of a 4 level attribute

+	00	01	10	11
00	00	01	10	11
01	01	00	11	10
10	10	11	00	01
11	11	10	01	00

$\times$	00	01	10	11
00	00	00	00	00
01	00	01	10	11
10	00	10	11	01
11	00	11	01	10

Table 1.21: Addition (a) and multiplication (b) tables for $GF[4]$ - pairs notation

We now demonstrate how these two representations are useful when adding and multiplying in $GF[4]$. When adding, the pairs representation becomes useful, as we can add each component modulo $s = 2$. For example

$$01 + 10 = 11$$

or

$$11 + 01 = 10.$$

When multiplying, the multiplicative representation is useful, since $\alpha^3 = 1$. Then we have

$$\begin{aligned} \alpha^2 \times \alpha^2 &= \alpha^4 \\ &= \alpha^3 \times \alpha \\ &= \alpha. \end{aligned}$$

□

1.A.4 Special Matrices

We can represent field operations in matrix notation. This representation will become very useful in Chapter 6 when we obtain new designs from existing designs. There are some special types of matrices that are particularly useful in this regard. One of these is a *circulant matrix*. Horn and Johnson [1985] define a circulant matrix to be an $n \times n$ matrix of the form

$$A = \begin{bmatrix} a_1 & a_2 & a_3 & \dots & a_{n-1} & a_n \\ a_n & a_1 & a_2 & \dots & a_{n-2} & a_{n-1} \\ a_{n-1} & a_n & a_1 & \dots & a_{n-3} & a_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ a_2 & a_3 & a_4 & \dots & a_n & a_1 \end{bmatrix}.$$

The authors also state that circulant matrices are commutative under matrix multiplication.

The other special matrix that will be used extensively is a *permutation matrix.* Horn and Johnson [1985] define a permutation matrix to be an $n \times n$ matrix P, where exactly one entry in each row and column is equal to 1, and all other entries are 0.

If we pre–multiply a matrix by a permutation matrix, we obtain a re–ordering of the rows of the matrix. If we post–multiply a matrix by a permutation matrix, we obtain a re–ordering of the columns of the matrix. Theorem 1.A.1 gives a useful property of permutation matrices that we use later in the thesis.

■ **THEOREM 1.A.1.**
For any $n \times n$ permutation matrix P, $P^T = P^{-1}$. □

We can combine these definitions to obtain a basic circulant permutation matrix. A *basic circulant permutation matrix* is a matrix that is both a circulant matrix and a permutation matrix. Let $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$. Then the circulant permutation matrix P_1, say, with first row $(0, 0, \ldots, 0, 1)$ acts on $\boldsymbol{v}$ to give $\boldsymbol{v}P_1 = (v_2, \ldots, v_n, v_1)$.

■ **EXAMPLE 1.A.6.**
Consider the integers modulo 3. Suppose that we have a vector containing the elements of $\mathbb{Z}_3$, $(0, 1, 2)$. Then we can post–multiply by permutation matrices to perform addition on this vector modulo 3. Then post–multiplication by the matrix

$$P_1 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

will add 1 to each element in the vector $(0, 1, 2)$ modulo 3, that is,

$$\begin{aligned} (0,1,2) \cdot P_1 &= (0,1,2) \cdot \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \\ &= (1,2,0). \end{aligned}$$

Similarly, we can define a basic circulant permutation matrix that represents the addition of 2 to each entry in the levels vector

$$P_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = P_1^2$$

For completeness, we use I_3 to represent the addition of 0 to the levels vector. □

1.B Standard Designs

In this section, we introduce the most common types of designed experiment. We will use many of these standard designs to construct designs for choice experiments. We begin by considering block designs.

1.B.1 Block Designs

In order to obtain estimates efficiently when conducting a choice experiment, we need to consider methods that will optimise the information obtained from each respondent. One way to do this is to create an efficient designed experiment. Street and Street [1986] define a *design* to be a finite set of items X, and a family of subsets B_i of X, $\mathcal{B} = \{B_i | i = I\}$. Here the finite set of items is the set of all possible items defined by the level combinations of the attributes. For choice experiments each subset is a choice set. The design of a choice experiment will dictate which choice sets will be presented to the respondents.

If we consider every possible combination of m items in a choice experiment, then the number of subsets used in the design, or in our case, the number of choice sets presented to the respondent, can become very large, even for a modest subset size. For this reason, it is useful to consider a design with only some of the combinations of items. Street and Street [1986] define an *incomplete design* to be a design where at least one block does not contain every item in the set X.

While using an incomplete design makes an experiment smaller, care must be taken to ensure that the same amount of information is collected about each item. It is also desirable to be able to estimate the merit of each item independently of other items. One type of incomplete design that achieves these two goals is a *Balanced Incomplete Block Design* (BIBD). Street and Street [1986] define a BIBD as a design in which all the blocks contain the same number of items, all items appear in the same number of blocks and each pair of items appear in the same number of blocks. We define b to be the number of blocks in the design, each block containing m items. There are $L = \prod_{q=1}^{k} \ell_q$ items in total, each of which appears in r blocks. The number of times that each pair appears in the same block is denoted by λ. We write this as $BIBD(L, b, r, m, \lambda)$. Bailey [2008] gives a good discussion on the construction of BIBDs.

■ **EXAMPLE 1.B.1.**
Suppose that we are conducting an experiment with 7 possible items. Table 1.22 gives a BIBD that could be used for this experiment. We notice that each block has 3 items, each item appears in three blocks, and each pair of items appears in the same block exactly once. This design is a BIBD with $b = 7$, $m = 3$, $L = 7$, $r = 3$ and $\lambda = 1$. □

1	2	4
2	3	5
3	4	6
4	5	0
5	6	1
6	0	2
0	1	3

Table 1.22: A $BIBD(7, 7, 3, 3, 1)$

1.B.2 Factorial Designs

In many experiments, including choice experiments, the items can be described by a number of features, also known as *factors* or *attributes*. When running such experiments, we are usually more interested in testing hypotheses about the attributes than testing hypotheses about the items themselves. The appropriate design for doing this is called a *factorial design*.

Street [2007] defines a factor to be any feature of the experimental units which may affect the response observed in the experiment. Each factor may take one of several values. These values are called the *levels* of the factor. If there are k factors and the q^{th} factor has ℓ_q levels, for $1 \leq q \leq k$, then we speak of an $\ell_1 \times \ell_2 \times \ldots \times \ell_k$ factorial design.

Notice that Street [2007] uses the term factor instead of attribute. These terms can be used interchangeably, however the convention for choice experiments is to use attribute, whereas in the theory of designed experiments the convention is to use factor. From here onwards we will use attribute, but we will still speak of a factorial design.

We can distinguish between a factorial design that includes all possible combinations of attribute levels, which is called a *complete factorial design*, and those that include a subset of these, called a *fractional factorial design*. In the next example, we show one type of each design for the same experiment.

■ **EXAMPLE 1.B.2.**
Consider an experiment with three attributes, each of which has three levels. Table 1.23(a) gives a full factorial design, and Table 1.23(b) gives a fractional factorial design for this experiment. Notice that the first design consists of all triples of three level attributes, and the second design contains only a third of these triples. □

(a)

Complete Factorial								
0	0	0	1	0	0	2	0	0
0	0	1	1	0	1	2	0	1
0	0	2	1	0	2	2	0	2
0	1	0	1	1	0	2	1	0
0	1	1	1	1	1	2	1	1
0	1	2	1	1	2	2	1	2
0	2	0	1	2	0	2	2	0
0	2	1	1	2	1	2	2	1
0	2	2	1	2	2	2	2	2

(b)

Fractional Factorial		
0	0	0
0	1	1
0	2	2
1	0	1
1	1	2
1	2	0
2	0	2
2	1	0
2	2	1

Table 1.23: A complete factorial design (a), and a fractional factorial design (b)

It is convenient to look at two different types of designs based on the number of levels each attribute can take. The first of these are designs where all of the attributes have the same number of levels. Such designs are called *symmetric designs*. The remaining designs have at least one pair of attributes that differ in the number of levels. Such designs are called *asymmetric designs*.

Contrasts

Now that we have designs for factorial experiments, we can look at how we can estimate the contribution of each attribute, and combinations of attribute levels to the attractiveness of each item, the goal of our experiment. We use main effects and interaction effects to do this.

In order to estimate main effects and interaction effects, we need to set up *contrasts*. Suppose that we order the items lexicographically, that is

$$(00\ldots00, 00\ldots01, \ldots, (\ell_1-1)(\ell_2-1)\ldots(\ell_k-1)).$$

Then a contrast is a linear function of the expected responses of the items such that the coefficients sum to 0. Consider a linear function $\sum_i \omega_i y_i$, where y_i is the response for item T_i, then

$$\tau = \mathcal{E}\left(\sum_i \omega_i y_i\right) = \sum_i \omega_i \mathcal{E}(y_i).$$

Such a linear function is said to be a contrast if $\sum_i \omega_i = 0$. If item T_i appears n_i times in the design then two contrasts with coefficients ζ_i and ω_i are said to be *orthogonal* if and only if

$$\sum_i \frac{\zeta_i \omega_i}{n_i} = 0.$$

Without loss of generality, we will scale the contrast coefficients so $\sum_i \omega_i^2 = 1$.

The *main effect* of an attribute is the effect of moving between levels of that attribute, averaged over the levels of the other attributes. Street [2007] defines the main effect of an attribute A in an ℓ^k factorial design as the comparison of responses among the ℓ sets $T_\theta = \{(x_1, x_2, \ldots, x_k)|x_a = \theta\}$, $\theta \in \{0, 1, \ldots, \ell-1\}$. The sets T_θ partition the ℓ^k items into ℓ sets each of size ℓ^{k-1}. We denote this partition $\mathcal{P}(A)$. We then have $\ell - 1$ contrasts that will form the basis of the main effect contrast subspace. In the next example, we give the partitions for the main effects for the full factorial design in Table 1.23.

■ **EXAMPLE 1.B.3.**

Consider the full factorial design in Table 1.23. We will label the three attributes A, B, and C. Then the main effect of attribute A will be constructed from the partitions based on the level of the first attribute

$$\begin{aligned}\mathcal{P}(A) = {} & \{\{000, 001, 002, 010, 011, 012, 020, 021, 022\}, \{100, 101, 102, 110, 111, 112, 120, 121, 122\}, \\ & \{200, 201, 202, 210, 211, 212, 220, 221, 222\}\}.\end{aligned}$$

Then the contrasts corresponding to the main effect of the first attribute will assign the same coefficient to each item in the same partition of $\mathcal{P}(A)$. The main effect of attribute B will be constructed using the partitions the based on the level of the second attribute

$$\begin{aligned}\mathcal{P}(B) = {} & \{\{000, 001, 002, 100, 101, 102, 200, 201, 202\}, \{010, 011, 012, 110, 111, 112, 210, 211, 212\}, \\ & \{020, 021, 022, 120, 121, 122, 220, 221, 222\}\}.\end{aligned}$$

Then the contrasts corresponding to the main effect of the second attribute will assign the same coefficient to each item in the same partition of $\mathcal{P}(B)$. The main effect of attribute C is defined similarly. □

While in many cases the estimation and testing of main effects are the primary consideration in an experiment, the interaction between attributes may also be of interest. For example, there may be little difference in preferences between a urine sample and a mouth swab when the sample is to be taken at home, but there may be a large difference if the sample is to be taken at the public clinic.

Street [2007] defines an *interaction effect* between attributes A and B, to be the comparison of the responses among the sets within the $\ell - 1$ partitions given by $\{(x_1, x_2, \ldots, x_k) | x_a + \theta x_b = \gamma\}$, $\theta, \gamma \in \{0, 1, \ldots, \ell_B - 1\}$, $\theta \neq 0$. For fixed θ, denote the partitions by $\mathcal{P}(AB^{\theta})$. There will be $(\ell_1 - 1)(\ell_2 - 1)$ orthogonal contrasts that will form the basis of the two–factor interaction contrast space. Higher order interactions are defined similarly. In the next example, we give the partitions for two–factor interactions for the full factorial design in Table 1.23.

■ **EXAMPLE 1.B.4.**
Once again, consider the full factorial design in Table 1.23, with three attributes A, B, and C. The coefficient of the interaction effect between attributes A and B can be based on contrasts between the sets of the partition, where such a contrast has the same coefficient for all items in the set. The contrast coefficient assigned depends on which partition of $\mathcal{P}(AB)$ the item belongs to, and which partition of $\mathcal{P}(AB^2)$ the item belongs to.

$$\begin{aligned}\mathcal{P}(AB) =& \{\{000, 001, 002, 120, 121, 122, 210, 211, 212\}, \{010, 011, 012, 100, 101, 102, 220, 221, 222\},\\ & \{020, 021, 022, 110, 111, 112, 200, 201, 202\}\}.\\ \mathcal{P}(AB^2) =& \{\{000, 001, 002, 110, 111, 112, 220, 221, 222\}, \{020, 021, 022, 100, 101, 102, 210, 211, 212\},\\ & \{010, 011, 012, 120, 121, 122, 200, 201, 202\}\}.\end{aligned}$$

This gives the partition

$$\{\{000, 001, 002\}, \{010, 011, 012\}, \{020, 021, 022\}, \{100, 101, 102\}, \{110, 111, 112\},\\ \{120, 121, 122\}, \{200, 201, 202\}, \{210, 211, 212\}, \{220, 221, 222\}\}.$$

Then the contrasts corresponding to the two–factor interactions between the first and second attributes will assign the same coefficient to each item in the same partition in the above partitioning so long as not all of the entries in the same partitions of $\mathcal{P}(A)$ have the same coefficients, and not all of the entries in the same partitions of $\mathcal{P}(B)$ have the same coefficients. The interaction between attributes A and C, and attributes B and C are defined similarly. □

We now look at an example of constructing a set of orthogonal contrasts corresponding to the main effects in a 2×3 factorial design.

■ **EXAMPLE 1.B.5.**
Consider a 2×3 factorial experiment. There are 6 possible items that we could present to the respondents. These are shown, in lexicographic order, in Table 1.24.

Suppose that we wish to compare the 0 and 1 levels of the first attribute. Then we could set the coefficients of the items with a 0 for the first attribute to be -1, and the coefficients of the items with a 1 for the first attribute to be 1. This gives

$$-\phi_{00} - \phi_{01} - \phi_{02} + \phi_{10} + \phi_{11} + \phi_{12}.$$

0	0
0	1
0	2
1	0
1	1
1	2

Table 1.24: The complete 2×3 factorial design

This satisfies the properties of a contrast, but not the additional scale requirement. If we divide each coefficient by $\sqrt{6}$, then $\sum \omega_i^2 = 1$. Then we have

$$-\frac{1}{\sqrt{6}}\phi_{00} - \frac{1}{\sqrt{6}}\phi_{01} - \frac{1}{\sqrt{6}}\phi_{02} + \frac{1}{\sqrt{6}}\phi_{10} + \frac{1}{\sqrt{6}}\phi_{11} + \frac{1}{\sqrt{6}}\phi_{12}.$$

We have one contrast for the first attribute since the first attribute has 2 levels, and $\ell - 1 = 1$ contrast will form a basis for the main effect subspace.

We can also define contrasts for the main effects of the second attribute. Since the second attribute has 3 levels, $\ell - 1 = 2$ orthogonal contrasts will form a basis for the main effect subspace. If we define one of these contrasts to have a coefficient of $-\frac{1}{2}$ for those items with the second attribute at level 0, a coefficient of 0 for those items with the second attribute at level 1, and a coefficient of $\frac{1}{2}$ for those items with the second attribute at level 2, then we obtain the contrast

$$-\frac{1}{2}\phi_{00} - 0 \times \phi_{01} + \frac{1}{2}\phi_{02} - \frac{1}{2}\phi_{10} + 0 \times \phi_{11} + \frac{1}{2}\phi_{12}.$$

Another contrast based on the three level attribute is

$$-\frac{1}{2\sqrt{3}}\phi_{00} + \frac{1}{\sqrt{3}}\phi_{01} - \frac{1}{2\sqrt{3}}\phi_{02} - \frac{1}{2\sqrt{3}}\phi_{10} + \frac{1}{\sqrt{3}}\phi_{11} - \frac{1}{2\sqrt{3}}\phi_{12}.$$

We can demonstrate that the two contrasts for the three level attribute are orthogonal to each other under equal replication (with $n_i = 1$ without loss of generality).

$$\begin{aligned}\sum_i \frac{\omega_i \zeta_i}{n_i} &= \frac{-1}{2} \times \frac{-1}{2\sqrt{3}} + 0 \times \frac{1}{\sqrt{3}} + \frac{1}{2} \times \frac{-1}{2\sqrt{3}} - \frac{1}{2} \times \frac{-1}{2\sqrt{3}} + 0 \times \frac{1}{\sqrt{3}} + \frac{1}{2} \times \frac{-1}{2\sqrt{3}} \\ &= 0,\end{aligned}$$

as required. □

We can write the coefficients for the contrasts as the rows of a matrix. We call such a matrix a *contrast matrix*, which we denote B. In a contrast matrix we label the columns using the items written in lexicographic order. The next example shows how a set of contrasts can be expressed as a contrast matrix

■ **EXAMPLE 1.B.6.**

We can place the contrast coefficients developed in Example 1.B.5 into a contrast matrix. This contrast matrix will be

$$B = \begin{array}{c} \begin{matrix} \scriptstyle 00 & \scriptstyle 01 & \scriptstyle 02 & \scriptstyle 10 & \scriptstyle 11 & \scriptstyle 12 \end{matrix} \\ \begin{bmatrix} \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{-1}{2} & 0 & \frac{1}{2} & \frac{-1}{2} & 0 & \frac{1}{2} \\ \frac{-1}{2\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{-1}{2\sqrt{3}} & \frac{-1}{2\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{-1}{2\sqrt{3}} \end{bmatrix} \end{array}.$$

□

In general, we assume that the contrasts in a contrast matrix have been chosen to be pairwise orthogonal, that is $BB^T = I$. When an attribute has ℓ levels, we need $\ell - 1$ contrasts to be the basis for the main effect subspace. There are many ways that these contrasts can be defined. The method that we will use in this thesis is to use orthogonal polynomial contrasts. Orthogonal polynomial contrasts are particularly useful when the attribute is continuous, since we are fitting a polynomial function that describes the change in response for a change in attribute level. Tables of these contrast coefficients can be found in texts such as Kuehl [2000].

While we can define contrasts to represent all main effects and higher order effects, we often do not wish to estimate them all. In this case, we partition the rows of B into two matrices, B_h and B_a. The contrasts whose coefficients are in B_h are those that we are interested in estimating, and the coefficients for the remaining contrasts will be contained in B_a. In general, we will be finding designs that are optimal for the estimation of the contrasts in B_h, but may not be able to estimate the contrasts in B_a at all.

■ **EXAMPLE 1.B.7.**

Consider again the 2×3 experiment in Example 1.B.5. The full set of orthogonal polynomial contrast coefficients for the estimation of main effects plus the two–factor interaction is

$$B = \begin{array}{c} \begin{matrix} \scriptstyle 00 & \scriptstyle 01 & \scriptstyle 02 & \scriptstyle 10 & \scriptstyle 11 & \scriptstyle 12 \end{matrix} \\ \begin{bmatrix} \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{-1}{2} & 0 & \frac{1}{2} & \frac{-1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{1}{2\sqrt{3}} & \frac{-1}{2\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{-1}{2\sqrt{3}} \\ \frac{1}{2} & 0 & \frac{-1}{2} & \frac{-1}{2} & 0 & \frac{1}{2} \\ \frac{-1}{2\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{-1}{2\sqrt{3}} & \frac{-1}{2\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{-1}{2\sqrt{3}} \end{bmatrix} \end{array}.$$

The first row of the matrix contains the contrast coefficients for the main effect of the first attribute. The next two rows contain the contrast coefficients for the main effect of the second attribute. The final two rows contain the contrast coefficients for the two–factor interaction.

Suppose that only the main effects are of interest. Then B_h will contain the coefficients corresponding to the main effects, and B_a will contain the coefficients corresponding to the two–factor interactions. That is,

$$B_h = \begin{bmatrix} \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{-1}{2} & 0 & \frac{1}{2} & \frac{-1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{1}{2\sqrt{3}} & \frac{-1}{2\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{-1}{2\sqrt{3}} \end{bmatrix},$$

and

$$B_a = \begin{bmatrix} \frac{1}{2} & 0 & \frac{-1}{2} & \frac{-1}{2} & 0 & \frac{1}{2} \\ \frac{-1}{2\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{-1}{2\sqrt{3}} & \frac{-1}{2\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{-1}{2\sqrt{3}} \end{bmatrix}.$$

□

The set of contrasts that we can use to form a basis for the main effect contrast space of an ℓ level attribute is not unique for $\ell \geq 4$. The next example illustrates this for $\ell = 4$.

■ **EXAMPLE 1.B.8.**

Consider a 4 level attribute. The orthogonal polynomial contrasts for a 4 level attribute are

$$B_{4(1)} = \begin{bmatrix} -\frac{3}{\sqrt{20}} & -\frac{1}{\sqrt{20}} & \frac{1}{\sqrt{20}} & \frac{3}{\sqrt{20}} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{\sqrt{20}} & \frac{3}{\sqrt{20}} & -\frac{3}{\sqrt{20}} & \frac{1}{\sqrt{20}} \end{bmatrix}.$$

Alteratively we could compare two pairs of levels, and then compare within each pair, giving the contrast matrix

$$B_{4(2)} = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

Finally, we could compare between three different pairings of the levels, giving

$$B_{4(3)} = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

These three contrast matrices contain different comparisons, but all form a basis for the main effects contrast space. □

1.B.3 Orthogonal Arrays

We now consider a design structure that is related to factorial designs, an *orthogonal array.* Street and Burgess [2007] define an orthogonal array $OA[N, k, \ell, t]$ to be an $N \times k$ array with elements from a set of ℓ symbols such that any $N \times t$ subarray has each t–tuple appearing as a row N/ℓ^t times. We say that t is the *strength* of the array, k is the number of *constraints* and ℓ is the number of *levels.*

Orthogonal arrays exist not only for symmetric designs, as in the definition above, but also for any combination of levels. Street and Burgess [2007] define an *asymmetric orthogonal array* $OA[N; \ell_1, \ell_2, \ldots, \ell_k; t]$ to be an $N \times k$ array with the elements in column q from a set of ℓ_q symbols such that any $N \times t$ subarray has each t–tuple appearing as a row an equal number of times.

Orthogonal arrays are useful in the design of choice experiments because they can be constructed in a number of ways, whilst retaining the properties that are important in designed experiments. One such feature is equal replication of each level of an attribute, which is assured when $t \geq 1$.

In this thesis, we will focus our attention on linear orthogonal arrays. Hedayat et al. [1999] defines an orthogonal array $OA[N, k, \ell, t]$ with levels taken from $GF[\ell]$ to be *linear* if all of its runs are distinct and if, when considered as k–tuples from $GF[\ell]$, its N runs form a vector space over $GF[\ell]$.

Rao [1947] and Rao [1949] give a method for constructing linear symmetric orthogonal arrays where the number of levels for each attribute is a prime power. This construction uses the properties of $GF[\ell]$, which we know to exist when ℓ is a prime power.

■ **CONSTRUCTION 1.B.1.**

(Rao–Hamming: Hedayat et al. [1999]) *Form an* $\ell^n \times n$ *array with all possible* n*–tuples from* $GF[\ell]$*. Let* $C_1, \ldots, C_n$ *denote the columns of this array. The columns of the full orthogonal array then consist of all columns of the form*

$$z_1 C_1 + z_2 C_2 + \ldots + z_n C_n = [C_1, C_2, \ldots, C_n]\boldsymbol{z}, \tag{1.1}$$

where $\boldsymbol{z} = (z_1, \ldots, z_n)^T$ *is an* n*–tuple from* $GF[\ell]$*, not all the* z_i *are zero, and the first non–zero* z_i *is* 1*. There are* $\frac{\ell^n - 1}{\ell - 1}$ *such columns.* □

In the next example, we illustrate this construction for a 3^4 experiment.

■ **EXAMPLE 1.B.9.**

In this example, we develop a 9 run OA with four 3–level attributes. We begin with all of the pairs of levels, as shown in the first two columns of Table 1.25. Then we need all pairs from $GF[3]$ where not all z_i are 0 and the first non–zero z_i is equal to 1. This gives the four pairs

$$(1, 0), (1, 1), (1, 2), \text{ and } (0, 1).$$

The first and last of these pairs, when substituted for the zs in Equation 1.1, will form the $\ell^n \times n$ array that we begin with. Then $\boldsymbol{z} = (1, 1)$ gives the column $C_1 + C_2$, and $\boldsymbol{z} = (1, 2)$ gives the column $C_1 + 2C_2$, where addition in both cases is in $GF[3]$. The resulting design is shown in Table 1.25. Observe that any 2 columns of this array contain every ordered pair exactly once. Thus $t = 2$. □

$\mathbf{C_1}$	$\mathbf{C_2}$	$\mathbf{C_1 + C_2}$	$\mathbf{C_1 + 2C_2}$
0	0	0	0
0	1	1	2
0	2	2	1
1	0	1	1
1	1	2	0
1	2	0	2
2	0	2	2
2	1	0	1
2	2	1	0

Table 1.25: The OA constructed in Example 1.B.9

This construction underlies the results obtained in Chapter 6.

1.B.4 Latin Squares

The last type of design that we consider in this section are Latin squares. Latin squares were originally used when both the horizontal and vertical positions of a plot of land were considered to be important. Subsequently they have been adopted for use in any situation where there are orthogonal blocking factors. Dénes and Keedwell [1974] define a *Latin square* to be an $n \times n$ matrix containing n different elements, each appearing in exactly once in each row and each column of the matrix. The integer n is called the *order* of the Latin square. An example of a Latin square of order 3 is

0	1	2
1	2	0
2	0	1

.

Dénes and Keedwell [1974] also show that the addition table for a finite group G is a Latin square of order ℓ.

We can sometimes impose an additional constraint on a Latin square to obtain a complete Latin square. Dénes and Keedwell [1974] define a *complete Latin square* to be a Latin square where for any ordered pair of distinct elements (a, b), with $1 \leq a, b \leq n$, there exists a row of the Latin square in which a appears to the right of b and a column of the Latin square in which a immediately precedes b. We use complete Latin squares in Chapter 5 to ensure that each item appears to the right of every other item exactly once in some choice set.

We conclude this section by looking at two constructions for complete Latin squares. The first of these was given in Williams [1949]. We call this the *Williams construction*.

■ **THEOREM 1.B.1.**

(Williams [1949])*Let $n = 2m$ be any even positive integer and let an $n \times n$ Latin square L be formed whose first row and column is $0, 1, 2m-1, 2, 2m-2, \ldots, m+1, m$, where the integers 0 to $2m-1$ are regarded as residues modulo n, and*

$$L_{i,j} = L_{i,1} + L_{1,j} \mod n,$$

then L is a complete Latin square. □

In the next example, we construct a complete Latin square of order 4 using this method.

■ **EXAMPLE 1.B.10.**

Suppose that we wish to construct a complete Latin square of order 4. Thus $n = 4$, and $m = 2$. Then by Theorem 1.B.1 the first row and column of the complete Latin square will be

$$[0 \quad 1 \quad 3 \quad 2].$$

Using Theorem 1.B.1 we obtain the complete Latin square in Table 1.26. □

0	1	3	2
1	2	0	3
3	0	2	1
2	3	1	0

Table 1.26: The complete Latin square constructed in Example 1.B.10

The sequence $0, 1, 2m-1, 2, \ldots, m-1, m$ is not the only sequence that yields a complete Latin square. Gordon [1961] shows that all partial orderings of a finite group yield a complete

Latin square. We define the *partial ordering* of a sequence $a_1, a_2, \ldots, a_n$ in a group of order n to be

$$a_1, a_1 \cdot a_2, a_1 \cdot a_2 \cdot a_3, \ldots, a_1 \cdot a_2 \cdot \ldots \cdot a_n.$$

This is best illustrated by an example.

■ **EXAMPLE 1.B.11.**
If $n = 6$ then

$$[0 \quad 5 \quad 2 \quad 3 \quad 4 \quad 1]$$

is a sequencing of $\mathbb{Z}_6$ with partial sums

$$[0 \quad 5 \quad 1 \quad 4 \quad 2 \quad 3].$$

By developing the row of partial sums in $\mathbb{Z}_6$ we obtain the complete Latin square in Table 1.27. □

0	5	1	4	2	3
5	4	0	3	1	2
1	0	2	5	3	4
4	3	5	2	0	1
2	1	3	0	4	5
3	2	4	1	5	0

Table 1.27: The complete Latin square constructed in Example 1.B.11

A list of partial sequencings of small order is given in Evans [2007].

Chapter 2

Choice Models that Incorporate Ties

Section 1.1 introduced the Davidson ties model as an extension of the Bradley–Terry model. In this chapter we look at optimal design theory for the Davidson ties model.

We start by looking at Davidson's original model, including deriving an expression for the determinant of the information matrix for a design, as defined in Section 1.3.1. We then find that from the set of competing designs used in Burgess and Street [2003], those designs that are optimal for the estimation of a set of attribute effects when the Bradley–Terry model is used are also optimal when the Davidson ties model is used to estimate the same set of attribute effects plus the ties parameter.

We will now consider an example to motivate our discussion of models incorporating ties.

■ **EXAMPLE 2.0.12.**
Consider a smaller version of the experiment presented in Example 1.0.1. This experiment has two attributes with two levels each. Table 2.1 gives the attributes and levels for this small experiment, as well as a coding for the levels of each attribute. Then there are 4 possible combinations of location and collection method:

- Draw blood at a public clinic (Coded 00),
- Draw blood at a doctor's office (Coded 01),
- Swab mouth at a public clinic (Coded 10), and
- Swab mouth at a doctor's office (Coded 11).

We wish to present choice sets to the respondent in such a way that if a respondent finds some of the options equally attractive, they may say so. In that case, we say those options have *tied*. □

We will return to this example as we progress through the chapter.

2.1 Review of the Davidson Ties Model

We begin by reviewing the results of Davidson [1970]. In particular we will recap some of the properties of the model that have already been developed in the literature, such as the distribution

Attribute	**Levels**	**Coding**
Sample collection	Draw blood	0
	Swab mouth/oral fluids	1
Location	Public clinic	0
	Doctor's office	1

Table 2.1: Attributes and levels for the HIV experiment with $k = 2$ and $\ell_1 = \ell_2 = 2$.

of the responses, the maximum likelihood estimates for the model, and the information matrix for the model. The logic used here will be used when the model is generalised in Chapter 3. We will use these results to show that the optimal designs presented in El-Helbawy and Ahmed [1984], Street et al. [2001] and Burgess and Street [2003] are also optimal when using the Davidson ties model.

Recall from Section 1.1 that the Davidson ties model is an extension to the Bradley–Terry model for paired comparisons. By using the Davidson ties model instead of the Bradley–Terry model we allow the respondent to state that they find the two items presented in the choice set equally attractive. This model captures this information through an additional parameter, $\nu > 0$, which measures how well the respondent can discriminate between the items in the choice set.

We have the following probabilities associated with each possible decision when the choice set $C = \{T_{i_1}, T_{i_2}\}$ is used.

$$P(T_{i_1}|C) = \frac{\pi_{i_1}}{\pi_{i_1} + \pi_{i_2} + \nu\sqrt{\pi_{i_1}\pi_{i_2}}},$$
$$P(T_{i_2}|C) = \frac{\pi_{i_2}}{\pi_{i_1} + \pi_{i_2} + \nu\sqrt{\pi_{i_1}\pi_{i_2}}},$$

and

$$P(T_{i_1} \text{ or } T_{i_2}|C) = \frac{\nu\sqrt{\pi_{i_1}\pi_{i_2}}}{\pi_{i_1} + \pi_{i_2} + \nu\sqrt{\pi_{i_1}\pi_{i_2}}}.$$

We denote $P(T_{i_1} \text{ or } T_{i_2}|C)$ as $P(\{T_{i_1}, T_{i_2}\}|C)$. Now we will see how these probabilities apply to our example.

■ **EXAMPLE 2.1.1.**

There are six possible choice sets of size 2 from the items listed in Example 2.0.12. If we consider the choice set $C = \{00, 11\}$ then the probability of choosing item 00 is

$$P(00|C) = \frac{\pi_{00}}{\pi_{00} + \pi_{11} + \nu\sqrt{\pi_{00}\pi_{11}}},$$

the probability of choosing item 11 is

$$P(11|C) = \frac{\pi_{11}}{\pi_{00} + \pi_{11} + \nu\sqrt{\pi_{00}\pi_{11}}},$$

and the probability of stating that the items are equally attractive is

$$P(\{00, 11\}|C) = \frac{\nu\sqrt{\pi_{00}\pi_{11}}}{\pi_{00} + \pi_{11} + \nu\sqrt{\pi_{00}\pi_{11}}}.$$

□

In his 1970 paper, Davidson derived the log–likelihood function and information matrix for this model. Because it will help us describe our generalisation, we provide a detailed derivation here. We will use the method presented here to derive the information matrix for the generalised Davidson ties model in Chapter 3.

Suppose that there are t items in total and that these are shown to the respondent in pairs. In each choice set the respondent may choose the item they prefer, or they can state that the two items presented are equally attractive. We define indicator variables w for subject α and task $C = \{T_{i_1}, T_{i_2}\}$ to represent the respondent's choice. We let

$$w_{\{i_1\}|C,\alpha} = \begin{cases} 1 & \text{if respondent } \alpha \text{ selects item } T_{i_1} \text{ when presented with the choice set } C, \\ 0 & \text{otherwise,} \end{cases}$$

$$w_{\{i_2\}|C,\alpha} = \begin{cases} 1 & \text{if respondent } \alpha \text{ selects item } T_{i_2} \text{ when presented with the choice set } C, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$w_{\{i_1,i_2\}|C,\alpha} = \begin{cases} 1 & \text{if respondent } \alpha \text{ finds items } T_{i_1} \text{ and } T_{i_2} \text{ equally} \\ & \text{attractive when presented with the choice set } C, \\ 0 & \text{otherwise,} \end{cases}$$

where $w_{\{i_1\}|C,\alpha} + w_{\{i_2\}|C,\alpha} + w_{\{i_1,i_2\}|C,\alpha} = 1$, since we do not allow repeated choice sets and do not have an opt-out process. For simplicity, we will proceed to write $w_{\{i\}|C,\alpha}$ as $w_{i|C,\alpha}$, but it useful to consider the outcome as the selection of a set with a single item. For a respondent α, the probability density function for their response to the choice set $C = \{T_{i_1}, T_{i_2}\}$ is

$$f_{C,\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \nu) = \frac{\pi_{i_1}^{w_{i_1|C,\alpha}} \pi_{i_2}^{w_{i_2|C,\alpha}} (\nu\sqrt{\pi_{i_1}\pi_{i_2}})^{w_{\{i_1,i_2\}|i_1 i_2 \alpha}}}{(\pi_{i_1} + \pi_{i_2} + \nu\sqrt{\pi_{i_1}\pi_{i_2}})^{n_C}},$$

where

$$\boldsymbol{w} = (w_{i_1|C,\alpha}, w_{i_2|C,\alpha}, w_{\{i_1,i_2\}|C,\alpha})^T,$$

and n_C is an indicator that equals 1 if the choice set C appears in the experiment and 0 if it does not. For consistency, if the choice set C does not appear in the experiment then we define

$$w_{i_1|C,\alpha} = w_{i_2|C,\alpha} = w_{\{i_1,i_2\}|C,\alpha} = 0.$$

The derivatives of the log of the density function with respect to each of the parameters are

$$\begin{aligned} \frac{\partial \ln(f_{C,\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \nu))}{\partial \pi_{i_1}} &= \frac{w_{i_1|C,\alpha}}{\pi_{i_1}} + \frac{w_{\{i_1,i_2\}|C,\alpha}}{2\pi_{i_1}} - \frac{n_C\left(1 + \frac{\pi_{i_2}\nu}{2\sqrt{\pi_{i_1}\pi_{i_2}}}\right)}{\pi_{i_1} + \pi_{i_2} + \nu\sqrt{\pi_{i_1}\pi_{i_2}}}, \\ \frac{\partial \ln(f_{C,\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \nu))}{\partial \pi_{i_2}} &= \frac{w_{i_2|C,\alpha}}{\pi_{i_2}} + \frac{w_{\{i_1,i_2\}|C,\alpha}}{2\pi_{i_2}} - \frac{n_C\left(1 + \frac{\pi_{i_1}\nu}{2\sqrt{\pi_{i_1}\pi_{i_2}}}\right)}{\pi_{i_1} + \pi_{i_2} + \nu\sqrt{\pi_{i_1}\pi_{i_2}}}, \\ \frac{\partial \ln(f_{C,\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \nu))}{\partial \pi_{i_3}} &= 0 \text{ where } i_3 \neq i_1, i_2, \end{aligned}$$

and

$$\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \nu))}{\partial \nu} = \frac{w_{\{i_1,i_2\}|C,\alpha}}{\nu} - \frac{n_C\sqrt{\pi_{i_1}\pi_{i_2}}}{\pi_{i_1} + \pi_{i_2} + \nu\sqrt{\pi_{i_1}\pi_{i_2}}}.$$

We will use these derivatives later to derive the entries of the information matrix for this model. We will now turn our attention to the maximum likelihood estimators for this model.

Since the likelihood function is the product of the density functions over all distinct choice sets and over all respondents, we have

$$\begin{aligned} L(\boldsymbol{w}, \boldsymbol{\pi}, \nu) &= \prod_{\alpha=1}^{s} \prod_{C} f_{C,\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \nu) \\ &= \prod_{C} \frac{\pi_{i_1}^{w_{i_1|C}} \pi_{i_2}^{w_{i_2|C}} (\nu\sqrt{\pi_{i_1}\pi_{i_2}})^{w_{\{i_1,i_2\}|C}}}{(\pi_{i_1} + \pi_{i_2} + \nu\sqrt{\pi_{i_1}\pi_{i_2}})^{sn_C}}, \end{aligned}$$

where $w_{i_1|C} = \sum_{\alpha=1}^{s} w_{i_1|C,\alpha}$, $w_{i_2|C} = \sum_{\alpha=1}^{s} w_{i_2|C,\alpha}$, and $w_{\{i_1,i_2\}|C} = \sum_{\alpha=1}^{s} w_{\{i_1,i_2\}|C,\alpha}$. Notice that n_C is not subscripted by respondent. This is because we will assume that all respondents are presented with the same set of choice sets.

To maximise this likelihood function, we need to set up a Lagrangian function to incorporate the restrictions placed on this model. For the purposes of convergence, we will enforce the normalising constraint present in the Bradley–Terry model

$$\sum_{i=1}^{t} \ln(\pi_i) = 0.$$

We will also constrain contrasts that we assume to be negligible to be equal to zero. If we let B_a be the matrix containing the coefficients of such contrasts, then we have

$$B_a \boldsymbol{\gamma} = \mathbf{0},$$

where $\boldsymbol{\gamma}$ is a vector containing $\gamma_i = \ln(\pi_i)$ for $i = 1, \ldots, t$. This will give the Lagrangian

$$\begin{aligned} G(\boldsymbol{w}, \boldsymbol{\pi}, \nu) &= \sum_{C} \Bigg(w_{i_1|C} \ln(\pi_{i_1}) + w_{i_2|C} \ln(\pi_{i_2}) + \frac{1}{2} w_{\{i_1,i_2\}|C} \big(\ln(\pi_{i_1}) + \ln(\pi_{i_2}) + 2\ln(\nu) \big) \\ &\quad - sn_C \ln(\pi_{i_1} + \pi_{i_2} + \nu\sqrt{\pi_{i_1}\pi_{i_2}}) \Bigg) + \kappa_1 \sum_{i=1}^{t} \ln(\pi_i) + [\kappa_2, \ldots, \kappa_{a+1}] B_a [\ln(\pi_i)], \end{aligned}$$

where κ_1 and κ_2 are Lagrange multipliers. If we differentiate $G(\boldsymbol{\pi}, \nu)$ with respect to π_i we obtain

$$\frac{\partial G(\boldsymbol{w}, \boldsymbol{\pi}, \nu)}{\partial \pi_i} = \sum_{C|T_i \in C} \left(\frac{w_{i|C}}{\pi_i} + \frac{w_{\{i,i_2\}|C}}{2\pi_i} - \frac{sn_C(1 + \frac{\pi_{i_2}\nu}{2\sqrt{\pi_i \pi_{i_2}}})}{\pi_i + \pi_{i_2} + \nu\sqrt{\pi_i \pi_{i_2}}} + \frac{\kappa_1}{\pi_i} \right) + \sum_{x=1}^{a} \kappa_{x+1} (B_a)_{xi} \frac{1}{\pi_i}, \tag{2.1}$$

and if we differentiate $G(\boldsymbol{\pi}, \nu)$ with respect to ν, we obtain

$$\frac{\partial G(\boldsymbol{w}, \boldsymbol{\pi}, \nu)}{\partial \nu} = \sum_{C} \left(\frac{w_{\{i_1,i_2\}|C}}{\nu} - \frac{sn_C\sqrt{\pi_{i_1}\pi_{i_2}}}{\pi_{i_1} + \pi_{i_2} + \nu\sqrt{\pi_{i_1}\pi_{i_2}}} \right).$$

We obtain the maximum likelihood estimates by setting these derivatives to zero and solving simultaneously. We can simplify this problem by using matrix notation. Suppose that we let

$$z_i = \sum_{C|T_i \in C} \left(w_{i|C} + \frac{1}{2} w_{\{i,i_2\}|C} - \frac{sn_C \hat{\pi}_i (1 + \frac{\hat{\pi}_{i_2}\hat{\nu}}{2\sqrt{\hat{\pi}_i \hat{\pi}_{i_2}}})}{\hat{\pi}_i + \hat{\pi}_{i_2} + \hat{\nu}\sqrt{\hat{\pi}_i \hat{\pi}_{i_2}}} \right).$$

Then by multiplying Equation 2.1 by π_{i_j} we get

$$z_i + \kappa_1 + \sum_{x=1}^{a} \kappa_{x+1} (B_a)_{xi} = 0.$$

This gives the system

$$\boldsymbol{z} + \kappa_1 \boldsymbol{j}_L + B_a^T \boldsymbol{\kappa} = \mathbf{0}_L, \tag{2.2}$$

where $\boldsymbol{z} = (z_1, z_2, \ldots, z_t)^T$ and $\boldsymbol{\kappa} = (\kappa_2, \ldots, \kappa_{a+1})^T$. Similarly, if we let

$$p = \sum_C \left(\frac{w_{\{i_1,i_2\}|C}}{\nu} - \frac{sn_C\sqrt{\pi_{i_1}\pi_{i_2}}}{\pi_{i_1} + \pi_{i_2} + \nu\sqrt{\pi_{i_1}\pi_{i_2}}} \right),$$

then we obtain $p = 0$ as the other constraint. If we pre–multiply Equation 2.2 by $\boldsymbol{j}_L^T$, we obtain

$$\kappa_1 = 0,$$

since $\boldsymbol{j}^T\boldsymbol{z}$ is shown to be zero in Appendix 2.A, and $\boldsymbol{j}^T B_a^T = \mathbf{0}$ since the rows of B_a are the coefficients of contrasts which are defined to add to zero.

We can pre–multiply Equation 2.2 by B_a to obtain

$$\boldsymbol{\kappa} = -B_a\boldsymbol{z}.$$

Substituting this back into Equation 2.2, we get

$$(I - B_a^T B_a)\boldsymbol{z} = \mathbf{0},$$

and

$$p = 0$$

as the normal equations. These can be solved iteratively to find the MLEs.

■ **EXAMPLE 2.1.2.**

Recall the experiment presented in Example 2.0.12. Suppose that we present the choice sets

$$\{\{00, 11\}, \{01, 10\}\}$$

to 50 respondents. Table 2.2 gives possible responses summarised over all respondents for this experiment. Then the likelihood function for this experiment is

$$L(\boldsymbol{w}, \boldsymbol{\pi}, \nu) = \frac{\pi_{00}^2\pi_{11}^{37}(\nu\sqrt{\pi_{00}\pi_{11}})^{11}}{(\pi_{00} + \pi_{11} + \nu\sqrt{\pi_{00}\pi_{11}})^{50}} \times \frac{\pi_{01}^6\pi_{10}^{28}(\nu\sqrt{\pi_{01}\pi_{10}})^{16}}{(\pi_{01} + \pi_{10} + \nu\sqrt{\pi_{01}\pi_{10}})^{50}}.$$

Option 1	Option 2	$\mathbf{T_1}$	$\mathbf{T_2}$	$\{\mathbf{T_1}, \mathbf{T_2}\}$
0 0	1 1	$w_{00\|C} = 2$	$w_{11\|C} = 37$	$w_{\{00,11\}\|C} = 11$
0 1	1 0	$w_{01\|C} = 6$	$w_{10\|C} = 28$	$w_{\{00,11\}\|C} = 16$

Table 2.2: Responses for the experiment in Example 2.1.2.

Now suppose that we are interested in the estimation of main effects and ν only. Then we assume that the two–factor interaction is negligible. This gives

$$B_a = \frac{1}{2}[1 \quad -1 \quad -1 \quad 1].$$

Then we have the constraints

$$\ln(\pi_{00}) + \ln(\pi_{01}) + \ln(\pi_{10}) + \ln(\pi_{11}) = 0,$$

and

$$\ln(\pi_{00}) - \ln(\pi_{01}) - \ln(\pi_{10}) + \ln(\pi_{11}) = 0.$$

This gives the Lagrangian

$$\begin{aligned}
&G(\boldsymbol{w},\boldsymbol{\pi},\nu)\\
&= 2\ln(\pi_{00}) + 37\ln(\pi_{11}) + \frac{11}{2}\ln(\pi_{00}) + \frac{11}{2}\ln(\pi_{11}) + 11\ln(\nu) - 50\ln(\pi_{00}+\pi_{11}+\nu\sqrt{\pi_{00}\pi_{11}})\\
&\quad +6\ln(\pi_{01}) + 28\ln(\pi_{10}) + \frac{16}{2}\ln(\pi_{01}) + \frac{16}{2}\ln(\pi_{10}) + 16\ln(\nu) - 50\ln(\pi_{01}+\pi_{10}+\nu\sqrt{\pi_{01}\pi_{10}})\\
&\quad +\kappa_1\left[\ln(\pi_{00}) + \ln(\pi_{01}) + \ln(\pi_{10}) + \ln(\pi_{11})\right] + \kappa_2\left[\ln(\pi_{00}) - \ln(\pi_{01}) - \ln(\pi_{10}) + \ln(\pi_{11})\right].
\end{aligned}$$

We differentiate $G(\boldsymbol{w},\boldsymbol{\pi},\nu)$ with respect to each π_i in turn to obtain

$$\begin{aligned}
\frac{\partial G(\boldsymbol{w},\boldsymbol{\pi},\nu)}{\partial \pi_{00}} &= \frac{15}{2\pi_{00}} - \frac{50(1+\frac{\nu\pi_{11}}{\sqrt{\pi_{00}\pi_{11}}})}{\pi_{00}+\pi_{11}+\nu\sqrt{\pi_{00}\pi_{11}}} + \frac{\kappa_1}{\pi_{00}} + \frac{\kappa_2}{\pi_{00}},\\
\frac{\partial G(\boldsymbol{w},\boldsymbol{\pi},\nu)}{\partial \pi_{01}} &= \frac{14}{\pi_{01}} - \frac{50(1+\frac{\nu\pi_{10}}{\sqrt{\pi_{01}\pi_{10}}})}{\pi_{01}+\pi_{10}+\nu\sqrt{\pi_{01}\pi_{10}}} + \frac{\kappa_1}{\pi_{01}} - \frac{\kappa_2}{\pi_{01}},\\
\frac{\partial G(\boldsymbol{w},\boldsymbol{\pi},\nu)}{\partial \pi_{10}} &= \frac{36}{\pi_{10}} - \frac{50(1+\frac{\nu\pi_{01}}{\sqrt{\pi_{01}\pi_{10}}})}{\pi_{01}+\pi_{10}+\nu\sqrt{\pi_{01}\pi_{10}}} + \frac{\kappa_1}{\pi_{10}} - \frac{\kappa_2}{\pi_{10}},
\end{aligned}$$

and

$$\frac{\partial G(\boldsymbol{w},\boldsymbol{\pi},\nu)}{\partial \pi_{11}} = \frac{85}{2\pi_{11}} - \frac{50(1+\frac{\nu\pi_{00}}{\sqrt{\pi_{00}\pi_{11}}})}{\pi_{00}+\pi_{11}+\nu\sqrt{\pi_{00}\pi_{11}}} + \frac{\kappa_1}{\pi_{11}} + \frac{\kappa_2}{\pi_{11}}.$$

Differentiating $G(\boldsymbol{w},\boldsymbol{\pi},\nu)$ with respect to ν gives

$$\frac{\partial G(\boldsymbol{w},\boldsymbol{\pi},\nu)}{\partial \nu} = \frac{27}{\nu} - \frac{50\sqrt{\pi_{00}\pi_{11}}}{(\pi_{00}+\pi_{11}+\nu\sqrt{\pi_{00}\pi_{11}})^2} - \frac{50\sqrt{\pi_{01}\pi_{10}}}{(\pi_{01}+\pi_{10}+\nu\sqrt{\pi_{01}\pi_{10}})^2}.$$

If we set each of these to 0 and solve iteratively we obtain the maximum likelihood estimates for $\boldsymbol{\pi}$ and ν. If we denote τ_1 as the main effect of the first attribute and τ_2 the main effect of the second attribute then we get

$$\hat{\tau_1} = 1.11 \qquad \hat{\tau_2} = 0.34 \qquad \hat{\nu} = 1.25. \qquad \square$$

2.2 Properties of the Davidson ties model

In this section we complete the construction of the information matrix for the estimation of the π_is and ν. We begin by deriving expressions for the expectations, variances and covariances of the selection indicators in $\boldsymbol{w}$. We then use these expressions to simplify the information matrix for $\boldsymbol{\pi}$ and ν.

Recall that $w_{i_1|C,\alpha}$, $w_{i_2|C,\alpha}$ and $w_{\{i_1 i_2\}|C,\alpha}$ are the selection indicators for the choice made by respondent α when presented with the choice set $C = \{T_{i_1}, T_{i_2}\}$. These indicators have a Bernoulli distribution with expectations

$$\begin{aligned}
\mathcal{E}_\pi(w_{i_1|C,\alpha}) &= \frac{\pi_{i_1}}{\pi_{i_1}+\pi_{i_2}+\nu\sqrt{\pi_{i_1}\pi_{i_2}}},\\
\mathcal{E}_\pi(w_{i_2|C,\alpha}) &= \frac{\pi_{i_2}}{\pi_{i_1}+\pi_{i_2}+\nu\sqrt{\pi_{i_1}\pi_{i_2}}},
\end{aligned}$$

and

$$\mathcal{E}_\pi(w_{\{i_1 i_2\}|C,\alpha}) = \frac{\nu\sqrt{\pi_{i_1}\pi_{i_2}}}{\pi_{i_1}+\pi_{i_2}+\nu\sqrt{\pi_{i_1}\pi_{i_2}}}. \qquad (2.3)$$

The variances of these indicators are

$$\begin{aligned}
\text{Var}_\pi(w_{i_1|C,\alpha}) &= \frac{\pi_{i_1}(\pi_{i_2}+\nu\sqrt{\pi_{i_1}\pi_{i_2}})}{(\pi_{i_1}+\pi_{i_2}+\nu\sqrt{\pi_{i_1}\pi_{i_2}})^2}, \\
\text{Var}_\pi(w_{i_2|C,\alpha}) &= \frac{\pi_{i_2}(\pi_{i_1}+\nu\sqrt{\pi_{i_1}\pi_{i_2}})}{(\pi_{i_1}+\pi_{i_2}+\nu\sqrt{\pi_{i_1}\pi_{i_2}})^2},
\end{aligned}$$

and

$$\text{Var}_\pi(w_{\{i_1i_2\}|C,\alpha}) = \frac{\nu\sqrt{\pi_{i_1}\pi_{i_2}}(\pi_{i_1}+\pi_{i_2})}{(\pi_{i_1}+\pi_{i_2}+\nu\sqrt{\pi_{i_1}\pi_{i_2}})^2}. \tag{2.4}$$

Next we derive the covariances between the selection indicators. First consider the covariance of two selection indicators for the selection of T_{i_1} from the choice set $C = \{T_{i_1}, T_{i_3}\}$ and the item T_{i_2} from the choice set $C' = \{T_{i_2}, T_{i_4}\}$, where $i_1 \neq i_2$. If the selections made in two distinct choice sets are uncorrelated, then

$$\begin{aligned}
\text{Cov}_\pi(w_{i_1|C,\alpha}, w_{i_2|C',\alpha}) &= \mathcal{E}_\pi\Big(\big(w_{i_1|C,\alpha}-\mathcal{E}_\pi(w_{i_1|C,\alpha})\big)\big(w_{i_2|C',\alpha}-\mathcal{E}_\pi(w_{i_2|C',\alpha})\big)\Big) \\
&= \begin{cases} \mathcal{E}_\pi\Big(\big(w_{i_1|C,\alpha}-\mathcal{E}_\pi(w_{i_1|C,\alpha})\big)\big(w_{i_2|C,\alpha}-\mathcal{E}_\pi(w_{i_2|C,\alpha})\big)\Big), & \text{if } C = C', \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}$$

If we expand this expectation and notice that only one outcome is possible, we see that

$$\mathcal{E}_\pi(w_{i_1|C,\alpha}w_{i_2|C,\alpha}) = 0,$$

and hence we obtain

$$\text{Cov}_\pi(w_{i_1|C,\alpha}, w_{i_2|C',\alpha}) = \begin{cases} \frac{-\pi_{i_1}\pi_{i_2}}{(\pi_{i_1}+\pi_{i_2}+\nu\sqrt{\pi_{i_1}\pi_{i_2}})^2}, & \text{if } C = C' \text{ and } i_1 \neq i_2, \\ \text{Var}_\pi(w_{i_1|C,\alpha}), & \text{if } C = C' \text{ and } i_1 = i_2, \\ 0, & \text{otherwise.} \end{cases} \tag{2.5}$$

Similarly, we obtain the covariance of the selection indicators for the selection of item T_{i_1} from $C = \{T_{i_1}, T_{i_3}\}$ and stating that the items in $C' = \{T_{i_2}, T_{i_4}\}$ are equally attractive.

$$\text{Cov}_\pi(w_{i_1|C,\alpha}, w_{\{i_2,i_4\}|C',\alpha}) = \begin{cases} \frac{-\pi_{i_1}\nu\sqrt{\pi_{i_1}\pi_{i_2}}}{(\pi_{i_1}+\pi_{i_2}+\nu\sqrt{\pi_{i_1}\pi_{i_2}})^2}, & \text{if } C = C', \\ 0, & \text{otherwise.} \end{cases}$$

Finally, the covariance for the selection indicators for stating that the items are equally attractive in both choice sets is given by

$$\text{Cov}_\pi(w_{\{i_1,i_3\}|C,\alpha}, w_{\{i_2,i_4\}|C',\alpha}) = \begin{cases} \text{Var}_\pi(w_{\{i_1,i_3\}|C,\alpha}), & \text{if } C = C', \\ 0, & \text{otherwise.} \end{cases}$$

We can now find the expectations, variances and covariances for the selection indicators in our example.

■ **EXAMPLE 2.2.1.**

Consider the experiment in Example 2.1.2. In particular, consider the first choice set, $C = \{00, 11\}$. The expected values for the selection indicators are given by

$$\begin{aligned}
\mathcal{E}_\pi(w_{00|C,\alpha}) &= \frac{\pi_{00}}{\pi_{00}+\pi_{11}+\nu\sqrt{\pi_{00}\pi_{11}}}, \\
\mathcal{E}_\pi(w_{11|C,\alpha}) &= \frac{\pi_{11}}{\pi_{00}+\pi_{11}+\nu\sqrt{\pi_{00}\pi_{11}}},
\end{aligned}$$

and

$$\mathcal{E}_\pi(w_{\{00,11\}|C,\alpha}) \quad = \quad \frac{\nu\sqrt{\pi_{00}\pi_{11}}}{\pi_{00}+\pi_{11}+\nu\sqrt{\pi_{00}\pi_{11}}}.$$

The variances of each choice are given by

$$\begin{aligned}
\text{Var}_\pi(w_{00|C,\alpha}) &= \frac{\pi_{00}(\pi_{11}+\nu\sqrt{\pi_{00}\pi_{11}})}{(\pi_{00}+\pi_{11}+\nu\sqrt{\pi_{00}\pi_{11}})^2}, \\
\text{Var}_\pi(w_{11|C,\alpha}) &= \frac{\pi_{11}(\pi_{00}+\nu\sqrt{\pi_{00}\pi_{11}})}{(\pi_{00}+\pi_{11}+\nu\sqrt{\pi_{00}\pi_{11}})^2},
\end{aligned}$$

and

$$\text{Var}_\pi(w_{\{00,11\}|C,\alpha}) \quad = \quad \frac{\nu\sqrt{\pi_{00}\pi_{11}}(\pi_{00}+\pi_{00})}{(\pi_{00}+\pi_{11}+\nu\sqrt{\pi_{00}\pi_{11}})^2}.$$

The covariances for each pair of selection indicators are given by

$$\begin{aligned}
\text{Cov}_\pi(w_{00|C,\alpha}, w_{11|C,\alpha}) &= \frac{-\pi_{00}\pi_{11}}{(\pi_{00}+\pi_{11}+\nu\sqrt{\pi_{00}\pi_{11}})^2}, \\
\text{Cov}_\pi(w_{00|C,\alpha}, w_{\{00,11\}|C,\alpha}) &= \frac{-\pi_{00}\nu\sqrt{\pi_{00}\pi_{11}}}{(\pi_{00}+\pi_{11}+\nu\sqrt{\pi_{00}\pi_{11}})^2},
\end{aligned}$$

and

$$\text{Cov}_\pi(w_{11|C,\alpha}, w_{\{00,11\}|C,\alpha}) \quad = \quad \frac{-\pi_{11}\nu\sqrt{\pi_{00}\pi_{11}}}{(\pi_{00}+\pi_{11}+\nu\sqrt{\pi_{00}\pi_{11}})^2}.$$

□

Next we construct the information matrix for the Davidson ties model. This is easier if we partition the information matrix into four blocks

$$I(\boldsymbol{\pi},\nu) = \begin{bmatrix} I_{\pi\pi}(\boldsymbol{\pi},\nu) & I_{\nu\pi}(\boldsymbol{\pi},\nu) \\ I_{\pi\nu}(\boldsymbol{\pi},\nu) & I_{\nu\nu}(\boldsymbol{\pi},\nu) \end{bmatrix}.$$

$I_{\pi\pi}(\boldsymbol{\pi},\nu)$ is a $t \times t$ matrix containing minus the expected value of the second derivatives with respect to two of the entries in $\boldsymbol{\pi}$. $I_{\pi\nu}(\boldsymbol{\pi},\nu)$ is a $t \times 1$ vector that contains minus the expected value of the second derivatives with respect to one entry in $\boldsymbol{\pi}$ and ν, and $I_{\nu\pi}(\boldsymbol{\pi},\nu) = I_{\pi\nu}(\boldsymbol{\pi},\nu)^T$. $I_{\nu\nu}(\boldsymbol{\pi},\nu)$ contains minus the expected value of the second derivative with respect to ν.

El-Helbawy and Bradley [1978] state that under some mild regularity conditions, as given in Section 1.1, the $(i,j)^{\text{th}}$ entry of the information matrix for a discrete choice experiment without ties is

$$I(\boldsymbol{\pi})_{ij} = \sum_{q=1}^{t-1}\sum_{r=q+1}^{t} \frac{n_{qr}}{N}\mathcal{E}_\pi\left(\left(\frac{\partial \ln(f_{qr\alpha}(\boldsymbol{\pi},\boldsymbol{w}))}{\partial \pi_i}\right)\left(\frac{\partial \ln(f_{qr\alpha}(\boldsymbol{\pi},\boldsymbol{w}))}{\partial \pi_j}\right)\right).$$

We now use this expression, and the results in Equations 2.3, 2.4, and 2.5, to evaluate some generic entries in each block matrix. To assist the generalisation in Chapter 3 we take the sum over all choice sets rather than the pairs of items, and modify the notation for n_{qr} and $f_{qr\alpha}(\boldsymbol{\pi},\boldsymbol{w})$ accordingly. We will begin with $I_{\pi\pi}(\boldsymbol{\pi},\nu)$. In this block matrix, we need to consider the off–diagonal and diagonal entries separately.

We begin with the generic off–diagonal entry. Consider $I_{\pi\pi}(\boldsymbol{\pi},\nu)_{ij}$, corresponding to product of the derivatives with respect to π_i and with respect to π_j,

$$I_{\pi\pi}(\boldsymbol{\pi},\nu)_{ij} \quad = \quad \sum_C \frac{n_C}{N}\mathcal{E}_\pi\left(\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{\pi},\boldsymbol{w},\nu))}{\partial \pi_i}\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{\pi},\boldsymbol{w},\nu))}{\partial \pi_j}\right).$$

We recall that the derivative of the density function for a particular choice set is zero if we differentiate with respect to an entry in $\boldsymbol{\pi}$ associated with an item that is not in the choice set. Thus both T_i and T_j must be in the choice set for the product to be non–zero. Since order does not matter when using this model, the choice sets $\{T_i, T_j\}$ and $\{T_j, T_i\}$ are equivalent. Then we can simplify to give

$$I_{\pi\pi}(\boldsymbol{\pi}, \nu)_{ij} = \frac{n_{\{i,j\}}}{N}\mathcal{E}_\pi\Bigg(\frac{\partial \ln(f_{\{i,j\},\alpha}(\boldsymbol{\pi}, \boldsymbol{w}, \nu))}{\partial \pi_i}\frac{\partial \ln(f_{\{i,j\},\alpha}(\boldsymbol{\pi}, \boldsymbol{w}, \nu))}{\partial \pi_j}\Bigg).$$

By observation, we obtain

$$\frac{\partial \ln(f_{\{i,j\},\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \nu))}{\partial \pi_i} = \frac{w_{i|\{i,j\},\alpha}}{\pi_i} + \frac{w_{\{i,j\}|\{i,j\},\alpha}}{2\pi_i} - \mathcal{E}_\pi\Big(\frac{w_{i|\{i,j\},\alpha}}{\pi_i} + \frac{w_{\{i,j\}|\{i,j\},\alpha}}{2\pi_i}\Big). \tag{2.6}$$

It then follows that

$$I_{\pi\pi}(\boldsymbol{\pi}, \nu)_{ij} = \frac{n_{\{i,j\}}}{N}\mathrm{Cov}_\pi\Big(\frac{w_{i|\{i,j\},\alpha}}{\pi_i} + \frac{w_{\{i,j\}|\{i,j\},\alpha}}{2\pi_i}, \frac{w_{j|\{i,j\},\alpha}}{\pi_j} + \frac{w_{\{i,j\}|\{i,j\},\alpha}}{2\pi_j}\Big).$$

We can now substitute Equation 2.5 to get

$$I_{\pi\pi}(\boldsymbol{\pi}, \nu)_{ij} = -\frac{n_{\{i,j\}}}{N}\frac{1}{4\pi_i\pi_j}\frac{4\pi_i\pi_j + (\pi_i + \pi_j)\nu\sqrt{\pi_i\pi_j}}{(\pi_i + \pi_j + \nu\sqrt{\pi_i\pi_j})^2}.$$

Next, we consider a generic diagonal entry of $I_{\pi\pi}(\boldsymbol{\pi}, \nu)$. This entry corresponds to the derivative with respect to π_i squared. We have

$$I_{\pi\pi}(\boldsymbol{\pi}, \nu)_{ii} = \sum_C \frac{n_C}{N}\mathcal{E}_\pi\Bigg(\Bigg(\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{\pi}, \boldsymbol{w}, \nu))}{\partial \pi_i}\Bigg)^2\Bigg).$$

Here, we notice that all choice sets which do not include T_i will have a derivative of 0, so we obtain

$$I_{\pi\pi}(\boldsymbol{\pi}, \nu)_{ii} = \sum_{C|T_i \in C} \frac{n_C}{N}\mathcal{E}_\pi\Bigg(\Bigg(\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{\pi}, \boldsymbol{w}, \nu))}{\partial \pi_i}\Bigg)^2\Bigg).$$

Using Equation 2.6 this becomes

$$I_{\pi\pi}(\boldsymbol{\pi}, \nu)_{ii} = \sum_{C|T_i \in C} \frac{n_C}{N}\mathrm{Var}_\pi\Bigg(\frac{w_{i|C,\alpha}}{\pi_i} + \frac{w_{\{i,i_2\}|C,\alpha}}{2\pi_i}\Bigg).$$

When we substitute Equations 2.4 and 2.5 we obtain

$$I_{\pi\pi}(\boldsymbol{\pi}, \nu)_{ii} = \sum_{C|T_i \in C} \frac{n_C}{4N\pi_i^2}\Bigg(\frac{4\pi_i\pi_{i_2} + \nu\sqrt{\pi_i\pi_{i_2}}(\pi_i + \pi_{i_2})}{(\pi_i + \pi_{i_2} + \nu\sqrt{\pi_i\pi_{i_2}})^2}\Bigg).$$

We now turn our attention to $I_{\pi\nu}(\boldsymbol{\pi}, \nu)$, a $t \times 1$ vector. We have

$$I_{\pi\nu}(\boldsymbol{\pi}, \nu)_i = \sum_C \frac{n_C}{N}\mathcal{E}_\pi\Bigg(\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{\pi}, \boldsymbol{w}, \nu))}{\partial \pi_i}\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{\pi}, \boldsymbol{w}, \nu))}{\partial \nu}\Bigg).$$

Notice that only the choice sets that include T_i will have a non–zero derivative. Then we obtain

$$I_{\pi\nu}(\boldsymbol{\pi}, \nu)_i = \sum_{C|T_i \in C} \frac{n_C}{N}\mathcal{E}_\pi\Bigg(\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{\pi}, \boldsymbol{w}, \nu))}{\partial \pi_i}\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{\pi}, \boldsymbol{w}, \nu))}{\partial \nu}\Bigg).$$

If we take into account Equation 2.6 and that

$$\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{\pi}, \boldsymbol{w}, \nu))}{\partial \nu} = \frac{w_{\{i_1,i_2\}|C,\alpha}}{\nu} - \mathcal{E}_\pi\left(\frac{w_{\{i_1,i_2\}|C,\alpha}}{\nu}\right) \tag{2.7}$$

then we obtain

$$I_{\pi\nu}(\boldsymbol{\pi},\nu)_i = \sum_{C|T_i\in C}\frac{n_C}{N}\mathrm{Cov}_\pi\left(\frac{w_{i|C,\alpha}}{\pi_i}+\frac{w_{\{i,i_2\}|C,\alpha}}{2\pi_i},\frac{w_{\{i,i_2\}|C,\alpha}}{\nu}\right).$$

We now substitute Equations 2.4 and 2.5 and simplify to obtain

$$I_{\pi\nu}(\boldsymbol{\pi},\nu)_i = \frac{1}{2N\pi_i\nu}\sum_{C|T_i\in C}\frac{n_C(\pi_{i_2}-\pi_i)\nu\sqrt{\pi_i\pi_{i_2}}}{(\pi_i+\pi_{i_2}+\nu\sqrt{\pi_i\pi_{i_2}})^2}.$$

These expressions will give the entries in the $I_{\pi\nu}(\boldsymbol{\pi},\nu)$ block matrix.

Finally, we look at the single element $I_{\nu\nu}(\boldsymbol{\pi},\nu)$. We begin with

$$I_{\nu\nu}(\boldsymbol{\pi},\nu) = \sum_C\frac{n_C}{N}\mathcal{E}_\pi\left(\left(\frac{\partial\ln(f_{C,\alpha}(\boldsymbol{\pi},\boldsymbol{w},\nu))}{\partial\nu}\right)^2\right).$$

If we use Equation 2.7, this simplifies to give

$$I_{\nu\nu}(\boldsymbol{\pi},\nu) = \sum_C\frac{n_C}{N}\mathrm{Var}_\pi\left(\frac{w_{\{i_1,i_2\}|C,\alpha}}{\nu}\right).$$

We then substitute Equation 2.4 and simplify, giving

$$I_{\nu\nu}(\boldsymbol{\pi},\nu) = \frac{1}{\nu N}\sum_C\frac{n_C(\pi_{i_1}+\pi_{i_2})\sqrt{\pi_{i_1}\pi_{i_2}}}{(\pi_{i_1}+\pi_{i_2}+\nu\sqrt{\pi_{i_1}\pi_{i_2}})^2}.$$

Since our ultimate goal is to test for main effects and interaction effects, which are linear combinations of the entries in $\boldsymbol{\gamma}=\ln(\boldsymbol{\pi})$, we need to first construct the information matrix for the estimation of the entries in $\boldsymbol{\gamma}$ and ν. This information matrix, introduced in Section 1.1, is denoted by $\Lambda(\boldsymbol{\pi},\nu)$. This takes the form

$$\Lambda(\boldsymbol{\pi},\nu) = PI(\boldsymbol{\pi},\nu)P^T,$$

where

$$P = \begin{bmatrix}\pi_1 & 0 & \dots & 0 & 0\\ 0 & \pi_2 & \dots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & 0 & \dots & \pi_t & 0\\ 0 & 0 & \dots & 0 & 1\end{bmatrix},$$

since $\frac{\partial\pi_i}{\partial\gamma_i}=\pi_i$ and $\frac{\partial\nu}{\partial\nu}=1$. Again it is convenient to partition the $\Lambda(\boldsymbol{\pi},\nu)$ matrix in the same way as we partitioned $I(\boldsymbol{\pi},\nu)$, giving

$$\Lambda(\boldsymbol{\pi},\nu) = \begin{bmatrix}\Lambda_{\gamma\gamma}(\boldsymbol{\pi},\nu) & \Lambda_{\nu\gamma}(\boldsymbol{\pi},\nu)\\ \Lambda_{\gamma\nu}(\boldsymbol{\pi},\nu) & \Lambda_{\nu\nu}(\boldsymbol{\pi},\nu)\end{bmatrix}.$$

Applying this to each of the generic entries in each block matrix and simplifying gives

$$\Lambda_{\gamma\gamma}(\boldsymbol{\pi},\nu)_{ii} = \sum_{C|T_i\in C}\frac{n_C}{4N}\left(\frac{4\pi_i\pi_{i_2}+\nu\sqrt{\pi_i\pi_{i_2}}(\pi_i+\pi_{i_2})}{(\pi_i+\pi_{i_2}+\nu\sqrt{\pi_i\pi_{i_2}})^2}\right),$$

$$\Lambda_{\gamma\gamma}(\boldsymbol{\pi},\nu)_{ij} = -\frac{n_{\{i,j\}}}{4N}\frac{4\pi_i\pi_j+(\pi_i+\pi_j)\nu\sqrt{\pi_i\pi_j}}{(\pi_i+\pi_j+\nu\sqrt{\pi_i\pi_j})^2},$$

$$\Lambda_{\gamma\nu}(\boldsymbol{\pi},\nu)_{1i} = \frac{1}{2N}\sum_{C|T_i\in C}\frac{n_C(\pi_{i_2}-\pi_i)\sqrt{\pi_i\pi_{i_2}}}{(\pi_i+\pi_{i_2}+\nu\sqrt{\pi_i\pi_{i_2}})^2},$$

and

$$\Lambda_{\nu\nu}(\boldsymbol{\pi}, \nu) = \frac{1}{N} \sum_C \frac{n_C(\pi_{i_1} + \pi_{i_2})\sqrt{\pi_{i_1}\pi_{i_2}}}{\nu(\pi_{i_1} + \pi_{i_2} + \nu\sqrt{\pi_{i_1}\pi_{i_2}})^2}. \quad (2.8)$$

If we make the assumption of equal merits then these entries simplify. We will leave ν unspecified as was done in Davidson [1970]. That is, we assume

$$\boldsymbol{\pi} = \mathbf{j} = \boldsymbol{\pi}_0,$$

and substitute into Equation 2.8 to obtain

$$\begin{aligned}
\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \nu)_{ii} &= \frac{1}{2N(2+\nu)} \sum_{C|T_i \in C} n_C, \\
\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \nu)_{ij} &= \frac{-n_{\{i,j\}}}{2N(2+\nu)}, \\
\Lambda_{\gamma\nu}(\boldsymbol{\pi}_0, \nu)_{1i} &= 0,
\end{aligned}$$

and

$$\Lambda_{\nu\nu}(\boldsymbol{\pi}_0, \nu) = \frac{2}{\nu(2+\nu)^2}.$$

These entries are defined for all $\nu > 0$, the range of possible values of ν for this model. If $\nu = 0$ then this model reduces to the Bradley–Terry model, and the $\Lambda(\boldsymbol{\pi}_0, \nu)$ matrix reduces to $\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \nu)$.

■ **EXAMPLE 2.2.2.**
Recall the experiment introduced in Example 2.0.12 and the design introduced in Example 2.1.2. The information matrix for the estimation of the entries in $\boldsymbol{\gamma}$ plus ν under the null hypothesis of equal merits is

$$\Lambda(\boldsymbol{\pi}_0, \nu) = \begin{bmatrix} \frac{1}{4(2+\nu)} & 0 & 0 & \frac{-1}{4(2+\nu)} & 0 \\ 0 & \frac{1}{4(2+\nu)} & \frac{-1}{4(2+\nu)} & 0 & 0 \\ 0 & \frac{-1}{4(2+\nu)} & \frac{1}{4(2+\nu)} & 0 & 0 \\ \frac{-1}{4(2+\nu)} & 0 & 0 & \frac{1}{4(2+\nu)} & 0 \\ 0 & 0 & 0 & 0 & \frac{2}{\nu(2+\nu)^2} \end{bmatrix}.$$

□

Now that we have a general expression for the information matrix for the estimation of the entries in $\boldsymbol{\gamma}$ plus ν, we can look at constructing information matrices for the estimation of contrasts of the entries in $\boldsymbol{\gamma}$ plus ν. We can then use these matrices to find designs that are optimal for the estimation of a set of contrasts of the entries in $\boldsymbol{\gamma}$ plus ν.

2.3 Representing options using k attributes

In this section we consider the construction of the information matrix when contrasts of the entries in $\boldsymbol{\gamma}$ and the estimation of ν are of interest. In particular, we are interested in contrasts of the entries in $\boldsymbol{\gamma}$ that represent the main effects of the attributes as introduced in Chapter 1.

Ideally, we would like to find the effect of level f_q of attribute q, denoted by β_{q,f_q}, or combinations of attribute levels on the merit of an item; that is, we want to estimate

$$\boldsymbol{\beta} = (\beta_{1,0}, \beta_{1,1}, \ldots, \beta_{1,\ell_1-1}, \ldots, \beta_{k,\ell_k-1}, \beta_{12,00}, \ldots, \beta_{12\ldots k,\ell_1-1\ldots\ell_k-1}, \nu)^T.$$

This is not possible however, because $\boldsymbol{\beta}$ is not estimable. It would be better to estimate contrasts of the entries in $\boldsymbol{\beta}$ so that we have a set of estimable contrasts. Suppose that the matrix B contains contrast coefficients that correspond to the coefficients of the effects that are of interest. We can choose the entries in B such that $B\boldsymbol{\beta}$ is estimable.

We now construct a matrix B_γ that contains coefficients of the contrasts of the entries in $\boldsymbol{\gamma}$. These contrasts may be the main effects of the attributes, or two–factor interactions between attributes, or perhaps subsets of these. We are not interested in the estimation of contrasts that include both ν and entries in $\boldsymbol{\gamma}$, but we do want to estimate ν itself. Then we assume that any interactions between $\boldsymbol{\gamma}$ and ν are not of interest, and therefore we assume that contrasts involving both entries in $\boldsymbol{\gamma}$ and ν are zero. Then we can construct a matrix B that contains both the contrast matrix B_γ and the effect of ν.

$$B = \begin{bmatrix} B_\gamma & \mathbf{0} \\ \mathbf{0} & B_\nu \end{bmatrix} = \begin{bmatrix} B_\gamma & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix}.$$

Then the information matrix for the estimation of the contrasts in B_γ and the ties parameter ν is

$$C(\boldsymbol{\pi}_0, \nu) = B\Lambda(\boldsymbol{\pi}_0, \nu)B^T,$$

which becomes

$$C(\boldsymbol{\pi}_0, \nu) = \begin{bmatrix} B_\gamma \Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \nu)B_\gamma^T & \mathbf{0} \\ \mathbf{0} & \frac{2}{\nu(2+\nu)^2} \end{bmatrix},$$

where $\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \nu)$ was defined in Section 2.2. Appendix 2.B shows that this information matrix does not violate the conditions given in El-Helbawy and Bradley [1978] that permit the transformation above. Since this information matrix is block diagonal, we are able to estimate the ties parameter independently of the attribute effects.

Now let us apply this to our example when we wish to estimate main effects and ν.

■ **EXAMPLE 2.3.1.**

Consider the experiment introduced in Example 2.0.12 and the design introduced in Example 2.1.2 for the estimation of main effects plus ν. The contrast matrix for the estimation of main effects plus ν is

$$B = \frac{1}{2}\begin{bmatrix} -1 & -1 & 1 & 1 & 0 \\ -1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix},$$

where B_γ is a 2×4 matrix of contrast coefficients and B_ν is the constant 1. Then the information matrix for the estimation of main effects plus ν is

$$C(\boldsymbol{\pi}_0, \nu) = \begin{bmatrix} \frac{1}{2(2+\nu)} & 0 & 0 \\ 0 & \frac{1}{2(2+\nu)} & 0 \\ 0 & 0 & \frac{2}{\nu(2+\nu)^2} \end{bmatrix}.$$

We see that we can estimate the main effects and ν independently when using this design, since the information matrix is diagonal. □

Now that we have an expression for the estimation of a set of contrasts that are of interest and the ties parameter we may now develop some results on the optimality of designs when using this model.

2.4 Optimal designs for the Davidson ties model

In this section we will compare the information matrices for the estimation of a set of effects when using the Bradley–Terry model and when using the Davidson ties model. Throughout this section, we assume that the same set of contrasts on the entries of $\boldsymbol{\gamma}$ are of interest, those whose coefficients are in B_γ. We will proceed to show that the optimal design for the estimation of a set of effects when estimating the Bradley–Terry model is also optimal for the estimation of the same set of effects plus ν when estimating the Davidson ties model.

Suppose that we assume $\boldsymbol{\pi} = \boldsymbol{\pi}_0$, then the information matrix for the estimation of the entries in $\boldsymbol{\gamma}$ when the Bradley–Terry model is used is denoted by $\Lambda(\boldsymbol{\pi}_0)_{\text{B–T}}$. Also suppose that $\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \nu)_{\text{DAV}}$ is the $(1,1)$ block of the information matrix for the estimation of the entries in $\boldsymbol{\gamma}$ plus ν when the Davidson ties model is used. Then if we compare the diagonal entries of both matrices, we see that

$$2(2+\nu)(\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \nu)_{\text{DAV}})_{ii} = 4(\Lambda(\boldsymbol{\pi}_0)_{\text{B–T}})_{ii} = \frac{1}{N}\sum_{C|T_i\in C} n_C.$$

If we compare the off–diagonal entries in the two matrices, we see that

$$2(2+\nu)(\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \nu)_{\text{DAV}})_{ij} = 4(\Lambda(\boldsymbol{\pi}_0)_{\text{B–T}})_{ij} = \frac{n_{\{i,j\}}}{N}.$$

It follows that we can express the information matrix for the estimation of $\boldsymbol{\gamma}$ and ν using the Davidson ties model in terms of the information matrix for the estimation of $\boldsymbol{\gamma}$ using the Bradley–Terry model. We get

$$\Lambda(\boldsymbol{\pi}_0, \nu)_{\text{DAV}} = \begin{bmatrix} \frac{2}{2+\nu}\Lambda(\boldsymbol{\pi}_0)_{\text{B–T}} & \mathbf{0} \\ \mathbf{0} & \frac{2}{\nu(2+\nu)^2} \end{bmatrix}.$$

Since we can express these information matrices in terms of each other, we may now look at comparing optimality results for these two designs. We will use the D–optimality criterion, as defined in Section 1.3.1.

■ **THEOREM 2.4.1.**
For a set of contrasts on the elements of $\boldsymbol{\gamma}$, a constant but unknown ν and for the same set of competing designs $\mathfrak{X}$, the D–optimal design for the estimation of the set of contrasts when using the Bradley–Terry model will also be D–optimal for the estimation of the same set of contrasts and ν when using the Davidson ties model under the null hypothesis of equal merits. □

Proof. We begin by letting B be the block diagonal matrix

$$B = \begin{bmatrix} B_\gamma & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix},$$

where B_γ is a $p \times t$ matrix containing the coefficients of the contrasts of interest among the entries in $\boldsymbol{\gamma}$. Then the information matrix for the estimation of $B_\gamma\boldsymbol{\gamma}$ plus ν when the Davidson ties model is used, $C(\boldsymbol{\pi}_0, \nu)_{\text{DAV}}$, is

$$\begin{aligned} C(\boldsymbol{\pi}_0, \nu)_{\text{DAV}} &= B\Lambda(\boldsymbol{\pi}_0, \nu)_{\text{DAV}}B^T \\ &= \begin{bmatrix} B_\gamma\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \nu)_{\text{DAV}}B_\gamma^T & \mathbf{0} \\ \mathbf{0} & \frac{2}{\nu(2+\nu)^2} \end{bmatrix}. \end{aligned}$$

Since we have shown that

$$\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \nu)_{\text{DAV}} = \frac{2}{2+\nu}\Lambda(\boldsymbol{\pi}_0)_{\text{B–T}},$$

by substitution we obtain

$$C(\boldsymbol{\pi}_0,\nu)_{\text{DAV}} = \begin{bmatrix} \frac{2}{2+\nu}B_\gamma\Lambda(\boldsymbol{\pi}_0)_{\text{B–T}}B_\gamma^T & \mathbf{0} \\ \mathbf{0} & \frac{2}{\nu(2+\nu)^2} \end{bmatrix}.$$

The information matrix for the estimation of $B_\gamma\boldsymbol{\gamma}$ when the Bradley–Terry model is used, $C(\boldsymbol{\pi}_0)_{\text{B–T}}$, is

$$C(\boldsymbol{\pi}_0)_{\text{B–T}} = B_\gamma\Lambda(\boldsymbol{\pi}_0)_{\text{B–T}}B_\gamma^T.$$

Thus we may express $C(\boldsymbol{\pi}_0,\nu)_{\text{DAV}}$ in terms of $C(\boldsymbol{\pi}_0)_{\text{B–T}}$ and ν, giving

$$C(\boldsymbol{\pi}_0,\nu)_{\text{DAV}} = \begin{bmatrix} \frac{2}{2+\nu}C(\boldsymbol{\pi}_0)_{\text{B–T}} & \mathbf{0} \\ \mathbf{0} & \frac{2}{\nu(2+\nu)^2} \end{bmatrix}.$$

Then we see that

$$\begin{aligned} \det(C(\boldsymbol{\pi}_0,\nu)_{DAV}) &= \det\left(\frac{2}{2+\nu}C(\boldsymbol{\pi}_0)_{\text{B–T}}\right) \times \frac{2}{\nu(2+\nu)^2} \\ &= \frac{2^{p+1}}{\nu(2+\nu)^{p+2}}\det(C(\boldsymbol{\pi}_0)_{\text{B–T}}). \end{aligned}$$

Since

$$\det((C(\boldsymbol{\pi}_0)_{\text{B–T}})_{\xi_{\text{OPT}}}) \geq \det((C(\boldsymbol{\pi}_0)_{\text{B–T}})_\xi)$$

for all $\xi \in \mathfrak{X}$, the relative efficiency of a generic design ξ compared to ξ_{OPT}, the design that is optimal for the estimation of $B_\gamma\boldsymbol{\gamma}$ when the Bradley–Terry model is used, when the Davidson ties model is used is

$$\begin{aligned} D_{\text{eff}}(\xi,\xi_{\text{OPT}}) &= \left(\frac{\det((C(\boldsymbol{\pi}_0,\nu)_{\text{DAV}})_\xi)}{\det((C(\boldsymbol{\pi}_0,\nu)_{\text{DAV}})_{\xi_{\text{OPT}}})}\right)^{\frac{1}{p+1}} \\ &= \left(\frac{\frac{2^{p+1}}{\nu(2+\nu)^{p+2}}\det((C(\boldsymbol{\pi}_0)_{\text{B–T}})_\xi)}{\frac{2^{p+1}}{\nu(2+\nu)^{p+2}}\det((C(\boldsymbol{\pi}_0)_{\text{B–T}})_{\xi_{\text{OPT}}})}\right)^{\frac{1}{p+1}} \\ &= \left(\frac{\det((C(\boldsymbol{\pi}_0)_{\text{B–T}})_\xi)}{\det((C(\boldsymbol{\pi}_0)_{\text{B–T}})_{\xi_{\text{OPT}}})}\right)^{\frac{1}{p+1}} \\ &\leq 1, \end{aligned}$$

for all $\xi \in \mathfrak{X}$. Therefore, by the definition of D–optimality, ξ_{OPT} is the D–optimal design for the estimation of the contrasts in B_γ plus ν when the Davidson ties model is used. □

We now consider an example of the relationship between these two models.

■ **EXAMPLE 2.4.1.**

Recall the experiment and design introduced in Example 2.1.2. In Example 2.3.1 we found the information matrix for the estimation of main effects plus ν when the Davidson ties model is used. Now we will find the information matrix for the estimation of main effects only using the same design and the Bradley–Terry model. The contrast matrix for the estimation of main effects is

$$B = \frac{1}{2}\begin{bmatrix} -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \end{bmatrix}.$$

From El-Helbawy and Bradley [1978] we know that, under the assumption of the null hypothesis of equal merits, the information matrix for the estimation of the entries in $\boldsymbol{\gamma}$ when the Bradley–Terry model is used is

$$\Lambda(\boldsymbol{\pi}_0, \nu)_{\text{B–T}} = \frac{1}{8}\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}.$$

We observe that

$$\Lambda_\gamma(\boldsymbol{\pi}_0, \nu)_{\text{DAV}} = \frac{2}{2+\nu}\Lambda(\boldsymbol{\pi}_0)_{\text{B–T}}.$$

It follows that the information matrix for the estimation of main effects only using the Bradley–Terry model is

$$C(\boldsymbol{\pi}_0)_{\text{B–T}} = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{4} \end{bmatrix},$$

and we see that

$$C(\boldsymbol{\pi}_0, \nu)_{\text{DAV}} = \begin{bmatrix} \frac{2}{2+\nu}C(\boldsymbol{\pi}_0)_{\text{B–T}} & 0 \\ 0 & \frac{2}{\nu(2+\nu)^2} \end{bmatrix}.$$

Taking determinants of both $C(\boldsymbol{\pi}_0, \nu)_{\text{DAV}}$ and $C(\boldsymbol{\pi}_0)_{\text{B–T}}$ gives

$$\det(C(\boldsymbol{\pi}_0, \nu)_{\text{DAV}}) = \frac{1}{2\nu(2+\nu)^4},$$

and

$$\det(C(\boldsymbol{\pi}_0)_{\text{B–T}}) = \frac{1}{16}.$$

We are estimating $p = 2$ contrasts on $\boldsymbol{\gamma}$, so

$$\begin{aligned} \frac{2^{p+1}}{\nu(2+\nu)^{p+2}}\det(C(\boldsymbol{\pi}_0)_{\text{B–T}}) &= \frac{2^3}{\nu(2+\nu)^4} \times \frac{1}{16} \\ &= \frac{1}{2\nu(2+\nu)^4} \\ &= \det(C(\boldsymbol{\pi}_0, \nu)_{\text{DAV}}), \end{aligned}$$

illustrating the results in Theorem 2.4.1. □

We now use this theorem to apply the known results for optimal designs when the Bradley–Terry model is used to the case of the Davidson ties model. First, we consider an extension of the theorem for a 2^k factorial experiment presented in El-Helbawy and Ahmed [1984].

■ **COROLLARY 2.4.2.**
Let ξ be the design that contains all distinct pairs that differ in the levels of each attribute in a 2^k paired comparisons experiment. Then when the rows of B_γ correspond to the k main effects, the design will be D–optimal for the estimation of the Davidson ties model. □

Proof. By Theorem 1.3.1, the design described in the statement of the theorem is D–optimal for the estimation of main effects when the Bradley–Terry model is used. It follows from Theorem 2.4.1 that this design must also be optimal for the estimation of the main effects plus the ties parameter when the Davidson ties model is used. □

We can use this corollary to find an optimal design for the estimation of main effects and ν for the experiment in our examples.

■ **EXAMPLE 2.4.2.**
Consider the 2^2 experiment introduced in Example 2.0.12. In this experiment we have $t = 4$ possible items. There are two pairs of items, $\{00, 11\}$ and $\{01, 10\}$, that differ in both attributes. Then the design with these two pairs is optimal for the estimation of main effects plus ν when the Davidson ties model is used. □

We can also extend the result on the optimal design for a 2^k factorial for the estimation of main effects plus two–factor interaction effects, established by Street et al. [2001], to incorporate ties.

■ **COROLLARY 2.4.3.**
The D–optimal design for testing main effects plus two–factor interactions and the ties parameter for a 2^k paired comparisons experiment, when all other effects are assumed zero, and the Davidson ties model is used is given by

$$a_{k,i} = \begin{cases} 2^{k-1}\binom{k}{(k+1)/2}^{-1}, & \text{if } i = \frac{k+1}{2}, \\ 0, & \text{otherwise,} \end{cases}$$

if k is odd, and

$$a_{k,i} = \begin{cases} 2^{k-1}\binom{k}{k/2}^{-1}, & \text{if } i = \frac{k}{2} \text{ or } \frac{k}{2}+1, \\ 0, & \text{otherwise,} \end{cases}$$

if k is even. □

Proof. By Theorem 1.3.2, the design described in the statement of the theorem is D–optimal for the estimation of main effects plus two–factor interactions when the Bradley–Terry model is used. It follows from Theorem 2.4.1 that this design must also be optimal for the estimation of main effects plus two–factor interactions and the ties parameter when the Davidson ties model is used. □

We can use this corollary to find an optimal design for the estimation of main effects plus two–factor interactions and ν for the experiment in our examples.

■ **EXAMPLE 2.4.3.**
Consider again the 2^2 experiment introduced in Example 2.0.12. We are now interested in the estimation of main effects plus two–factor interactions and ν. Since $k = 2$ is even the D–optimal design for the estimation of these effects is given by

$$a_{2,i} = \begin{cases} 2 \times \binom{2}{1}^{-1}, & \text{if } i = 1 \text{ or } 2, \\ 0, & \text{otherwise.} \end{cases}$$

This is the design with all pairs of distinct items. □

Finally, we can extend the results of the general factorial as presented in El-Helbawy et al. [1994] to incorporate ties.

■ **COROLLARY 2.4.4.**
Consider an $\ell_1 \times \ldots \times \ell_k$ factorial paired comparisons experiment. Assuming that there are no interactions present, and B_h consists of the main effects, then the design consisting of all pairs where the options differ in all of the attributes will be D–optimal in the design space for the estimation of main effects plus ν when the Davidson ties model is used. □

Proof. By Theorem 1.3.3, the design described in the statement of the theorem is D–optimal for the estimation of main effects when the Bradley–Terry model is used. It follows from Theorem 2.4.1 that this design must also be optimal for the estimation of main effects plus the ties parameter when the Davidson ties model is used. □

We now consider an example of how this result can be used to find optimal designs for the Davidson ties model.

■ **EXAMPLE 2.4.4.**
Let us consider the 3^2 experiment whose attributes and levels are given in Table 2.3. This experiment has 9 possible items. The optimal design for the estimation of main effects plus ν is given in Table 2.4. □

Attributes	**Levels**	**Coding**
Sample collection	Draw blood	0
	Swab mouth/oral fluids	1
	Urine sample	2
Location	Public clinic	0
	Doctor's office	1
	Home	2

Table 2.3: Attributes and levels for the HIV experiment with $\ell_1 = \ell_2 = 3$.

Option 1	**Option 2**	**Option 1**	**Option 2**
0 0	1 1	1 1	2 2
0 0	1 2	1 2	2 0
0 1	1 0	1 2	2 1
0 1	1 2	2 0	0 1
0 2	1 0	2 0	0 2
0 2	1 1	2 1	0 0
1 0	2 1	2 1	0 2
1 0	2 2	2 2	0 0
1 1	2 0	2 2	0 1

Table 2.4: Optimal design for the estimation of main effects and ν when $\ell_1 = \ell_2 = 3$.

2.5 Simulations of the Davidson ties model

In this section we consider the performance of the Davidson ties model under various model assumptions by carrying out a number of simulation studies. The simulations we perform here, and in later chapters, are based on a Type I extreme value error distribution. A comprehensive discussion of the simulation methods that we utilise here is given in Train [2003] page 209–210.

We assume that $k = 2$, $\ell_1 = \ell_2 = 2$ and $m = 2$ throughout. We consider two sets of values for the parameters. In the first we assume that both main effects parameters, τ_1 and τ_2, are equal to 0 and the ties parameter $\nu = 0.5$, and in the second set we assume that $\tau_1 = 1$ and $\tau_2 = -1$ but $\nu = 0.5$ still.

We find the locally optimal design for each set of values and compare the performance of each design with both sets of parameter values. The design in Table 2.5 is optimal for the estimation of the main effects of the attributes plus the ties parameter when $\tau_1 = \tau_2 = 0$, and $\nu = 0.5$, by Corollary 2.4.2. By an exhaustive search of the $2^6 - 1 = 63$ possible designs, we can show that the design in Table 2.6 is one of the designs that is optimal for the estimation of the main effects of the attributes plus the ties parameter when $\tau_1 = 1$, $\tau_2 = -1$, and $\nu = 0.5$. The other optimal design consists of the choice sets in Table 2.5 plus the choice set $\{\{1, 0\}, \{1, 1\}\}$. We will use the design in Table 2.6 for the simulations. The exhaustive search is illustrated in Figure 2.1, where the x–coordinate corresponds to the design index, and the y–coordinate is the determinant of the information matrix for that design when $\tau_1 = 1$, $\tau_2 = -1$, and $\nu = 0.5$. The determinants of the information matrix for the designs in Tables 2.5 and 2.6 are labelled in Figure 2.1.

We first assume that $\tau_1 = \tau_2 = 0$, and $\nu = 0.5$ and compare the simulated distributions of the parameter estimates when the designs in Tables 2.5 and 2.6 are used in turn. Each simulation is modelled using the simulated responses from 150 respondents, and each boxplot displays the distribution of the estimates from 1000 such simulations. Figures 2.2(a) and (b) show the distributions of the parameter estimates when the designs in Tables 2.5 and 2.6, respectively, are used. Summary statistics for both simulations are provided in Table 2.7. We see that, for both designs, the distribution of the parameter estimates seem to be unbiased and symmetric. We see that, in this case, the additional choice set in the design in Table 2.6 does not seem to reduce the variance of the τ_1 or ν, but the variance of τ_2 is reduced. This is reasonable given that the additional choice set requires the respondent to choose between the levels of the second attribute, while fixing the first attribute at level 0.

We now consider the performance of these two designs when $\tau_1 = 1, \tau_2 = -1$, and $\nu = 0.5$. Figures 2.3(a) and (b) show the distributions of the parameter estimates when the designs in Tables 2.5 and 2.6, respectively, are used. Summary statistics for both simulations are provided in Table 2.8. We see that, for both designs, the distribution of the parameter estimates seem to be unbiased and close to symmetric. For these parameter estimates, we see that the addition of an extra choice set does seem to reduce the variance of the parameter estimates across the board, but most markedly for τ_2. The selection probabilities when the design in Table 2.5 is used and $\tau_1 = 1, \tau_2 = -1$, and $\nu = 0.5$ are given in Table 2.9.

Next, we simulate the effect of changing the magnitude of the ties parameter on the distri-

Option 1	**Option 2**
0 0	1 1
0 1	1 0

Table 2.5: Optimal design for the estimation of main effects and ν when $\tau_1 = \tau_2 = 0$, and $\nu = 0.5$.

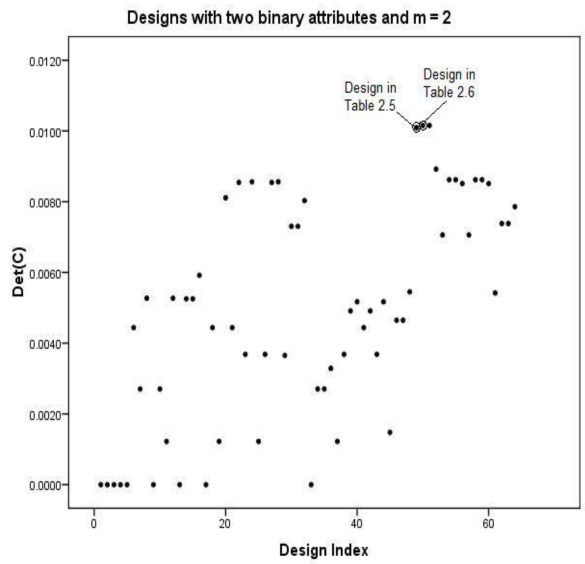

Figure 2.1: Exhaustive search for optimal design $\tau_1 = 1, \tau_2 = -1$, and $\nu = 0.5$.

Option 1	Option 2
0 0	1 1
0 1	1 0
0 0	0 1

Table 2.6: Optimal design for the estimation of main effects and ν when $\tau_1 = 1$, $\tau_2 = -1$, and $\nu = 0.5$.

Parameter	Simulated Mean (Standard Error)	Theoretical Variance	Simulated Variance	Simulated Skewness (Standard Error)
Design in Table 2.5				
τ_1	0.00308(0.00208)	0.00417	0.00434	0.03812(0.07734)
τ_2	−0.00002(0.00212)	0.00417	0.00451	−0.05835(0.07734)
ν	0.50383(0.00229)	0.00521	0.00525	0.24420(0.07734)
Design in Table 2.6				
τ_1	−0.00101(0.00201)	0.00417	0.00404	−0.14353(0.07734)
τ_2	0.00056(0.00166)	0.00278	0.00274	−0.09451(0.07734)
ν	0.50239(0.00184)	0.00347	0.00338	0.18347(0.07734)

Table 2.7: Summary statistics for $\tau_1 = \tau_2 = 0$, and $\nu = 0.5$.

(a) **(b)**

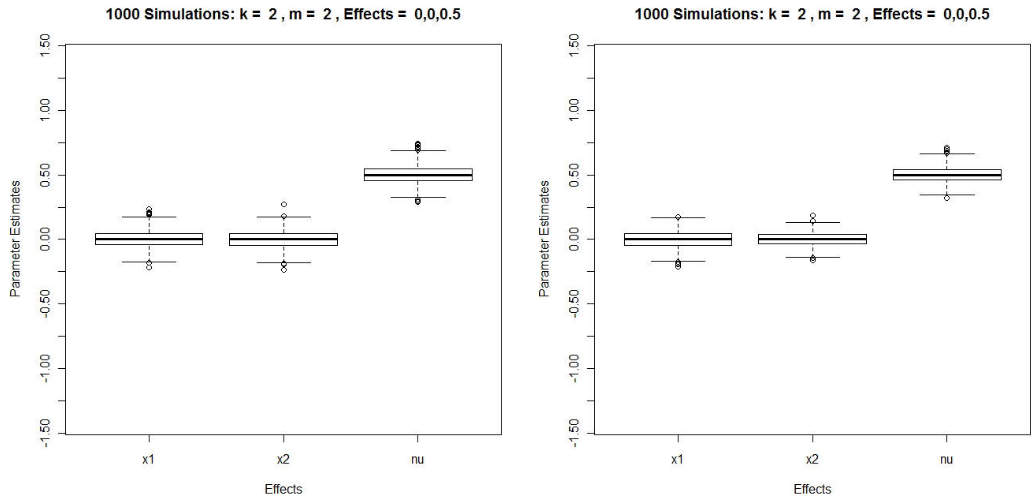

Figure 2.2: Simulation of Davidson ties model $\tau_1 = \tau_2 = 0$, and $\nu = 0.5$.

(a) **(b)**

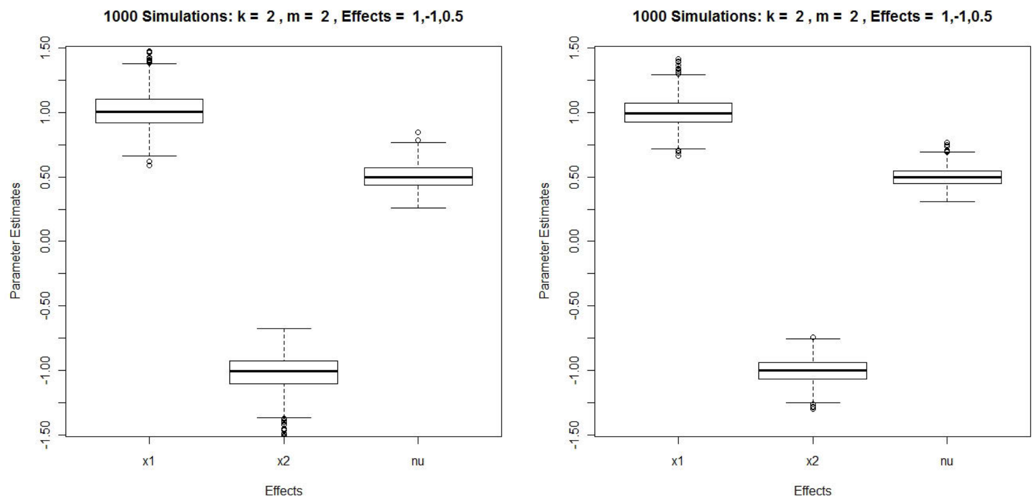

Figure 2.3: Simulation of Davidson ties model $\tau_1 = 1$, $\tau_2 = -1$, and $\nu = 0.5$.

Parameter	Simulated Mean (Standard Error)	Theoretical Variance	Simulated Variance	Simulated Skewness (Standard Error)
Design in Table 2.5				
τ_1	1.02064(0.00459)	0.00718	0.02103	0.48166(0.07734)
τ_2	−1.02260(0.00456)	0.00718	0.02077	−0.51109(0.07734)
ν	0.50506(0.00296)	0.00565	0.00878	0.24441(0.07734)
Design in Table 2.6				
τ_1	1.00382(0.00364)	0.00479	0.01327	0.33741(0.07734)
τ_2	−1.00187(0.00300)	0.00479	0.00902	−0.17943(0.07734)
ν	0.49988(0.00234)	0.00377	0.00550	0.22402(0.07734)

Table 2.8: Summary statistics for $\tau_1 = 1, \tau_2 = -1$, and $\nu = 0.5$.

Choice Set	$\mathbf{P(T_1 \mid \{T_1, T_2\})}$	$\mathbf{P(T_2 \mid \{T_1, T_2\})}$	$\mathbf{P(\{T_1, T_2\} \mid \{T_1, T_2\})}$
$\{00, 11\}$	0.400	0.400	0.200
$\{01, 10\}$	0.017	0.921	0.062

Table 2.9: Selection probabilities when $\tau_1 = 1$, $\tau_2 = -1$, and $\nu = 0.5$.

bution of the parameter estimates when we let $\tau_1 = 1$ and $\tau_2 = -1$, and use the design in Table 2.6. Figures 2.4 (a) and (b) give the simulated distributions of the parameter estimates when $\nu = 0.25$, and when $\nu = 1$, respectively. Summary statistics for both simulations are provided in Table 2.10. We see that the estimates are unbiased, and the variance of the estimates for τ_1 and τ_2 are similar for both simulations. The simulated variances of the parameter estimates for ν seem to be larger when $\nu = 1$. This is confirmed by looking at the theoretical variances of the parameter estimates, which also increase. We also notice that the distribution of ν is slightly right skewed. This is consistent with the assumption made in Critchlow and Fligner [1991] that $\ln(\nu)$ is normally distributed.

We now compare the ability of four different designs to estimate the main effects plus the two–factor interaction of the attributes and ν. The first two designs are those in Tables 2.5 and 2.6. The third design is the set of all pairs of items, which is optimal for the estimation of the main effects plus the two–factor interaction of the attributes and ν when $\tau_1 = \tau_2 = \tau_{12} = 0$, and $\nu = 0.5$, by Corollary 2.4.3. This design is shown in Table 2.11. The final design, shown in Table 2.12, is locally optimal for the estimation of the main effects plus two–factor interaction of the attributes and ν when $\tau_1 = 1, \tau_2 = -1, \tau_{12} = -0.25$, and $\nu = 0.5$, by an exhaustive search.

We first consider the case where the interaction effect is assumed to be negligible. Suppose that $\tau_1 = 1, \tau_2 = -1, \tau_{12} = 0$, and $\nu = 0.5$. Then Figures 2.5(a), (b), (c), and (d) give the simulated distributions of the parameter estimates when the designs in Table 2.5, Table 2.6, Table 2.11, and Table 2.12 are used. Summary statistics for all four of the simulations are provided in Table 2.13.

We see that the design in Table 2.5 can not be used to estimate the two–factor interaction

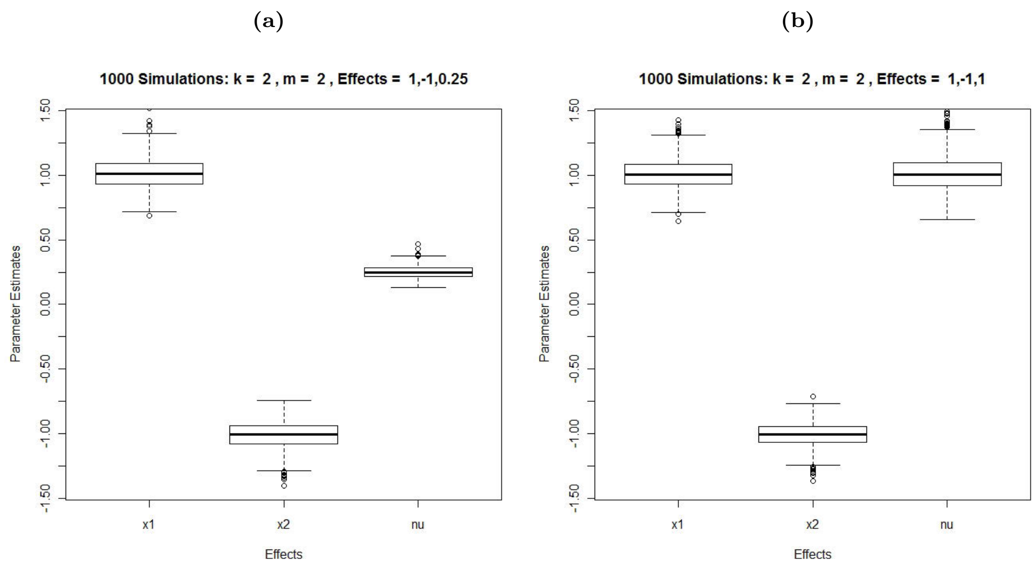

Figure 2.4: Simulation of Davidson ties model $\tau_1 = 1$, $\tau_2 = -1$, and (a) $\nu = 0.25$ (b) $\nu = 1$.

	Parameter	Simulated Mean (Standard Error)	Theoretical Variance	Simulated Variance	Simulated Skewness (Standard Error)
(a)	τ_1	1.01526(0.00376)	0.00929	0.01410	0.29054(0.07734)
	τ_2	−1.01345(0.00326)	0.00444	0.01065	−0.39184(0.07734)
	ν	0.25088(0.00151)	0.00223	0.00228	0.30937(0.07734)
(b)	τ_1	1.01229(0.00368)	0.01060	0.01358	0.32162(0.07734)
	τ_2	−1.01078(0.00313)	0.00558	0.00980	−0.31087(0.07734)
	ν	1.01487(0.00434)	0.01594	0.01881	0.48877(0.07734)

Table 2.10: Summary statistics for $\tau_1 = 1, \tau_2 = -1$, and (a) $\nu = 0.25$ and (b) $\nu = 1$.

Option 1	Option 2
0 0	0 1
0 1	1 1
0 1	1 0
1 0	0 0
1 0	1 1
1 1	0 0

Table 2.11: Optimal design for the estimation of main effects, two–factor interactions, and ν when $\tau_1 = \tau_2 = \tau_{12} = 0$, and $\nu = 0.5$.

Option 1	**Option 2**
0 0	0 1
0 1	1 1
1 0	0 0
1 0	1 1
1 1	0 0

Table 2.12: Optimal design for the estimation of main effects, two–factor interactions, and ν when $\tau_1 = 1, \tau_2 = -1, \tau_{12} = -0.25$, and $\nu = 0.5$.

at all, and gives biased estimates for the remaining attribute effects. The design in Table 2.6 is able to estimate the two–factor interaction, but with a relatively large variance and skewness toward the extremes. The designs in Tables 2.11 and 2.12 both give unbiased and symmetric parameter estimates with relatively small variances. The variance of the parameter estimates from the design in Table 2.11 is slightly lower than those from the design in Table 2.12.

Next, we consider the case where there is a non–zero interaction effect. Suppose that $\tau_1 = 1$, $\tau_2 = -1$, $\tau_{12} = -0.25$, and $\nu = 0.5$. Then Figures 2.6(a), (b), (c), and (d) give the simulated parameter estimates when the designs in Table 2.5, Table 2.6, Table 2.11, and Table 2.12 are used. Summary statistics for all four of the simulations are provided in Table 2.14.

Again we notice that the designs in Tables 2.11 and 2.12 give unbiased and reasonably symmetric parameter estimates with a relatively small variance. We also see that the designs in Tables 2.5 and 2.6 again give poorer estimates for these parameters. Once again, the design in Table 2.5 can not be used to estimate the two–factor interaction.

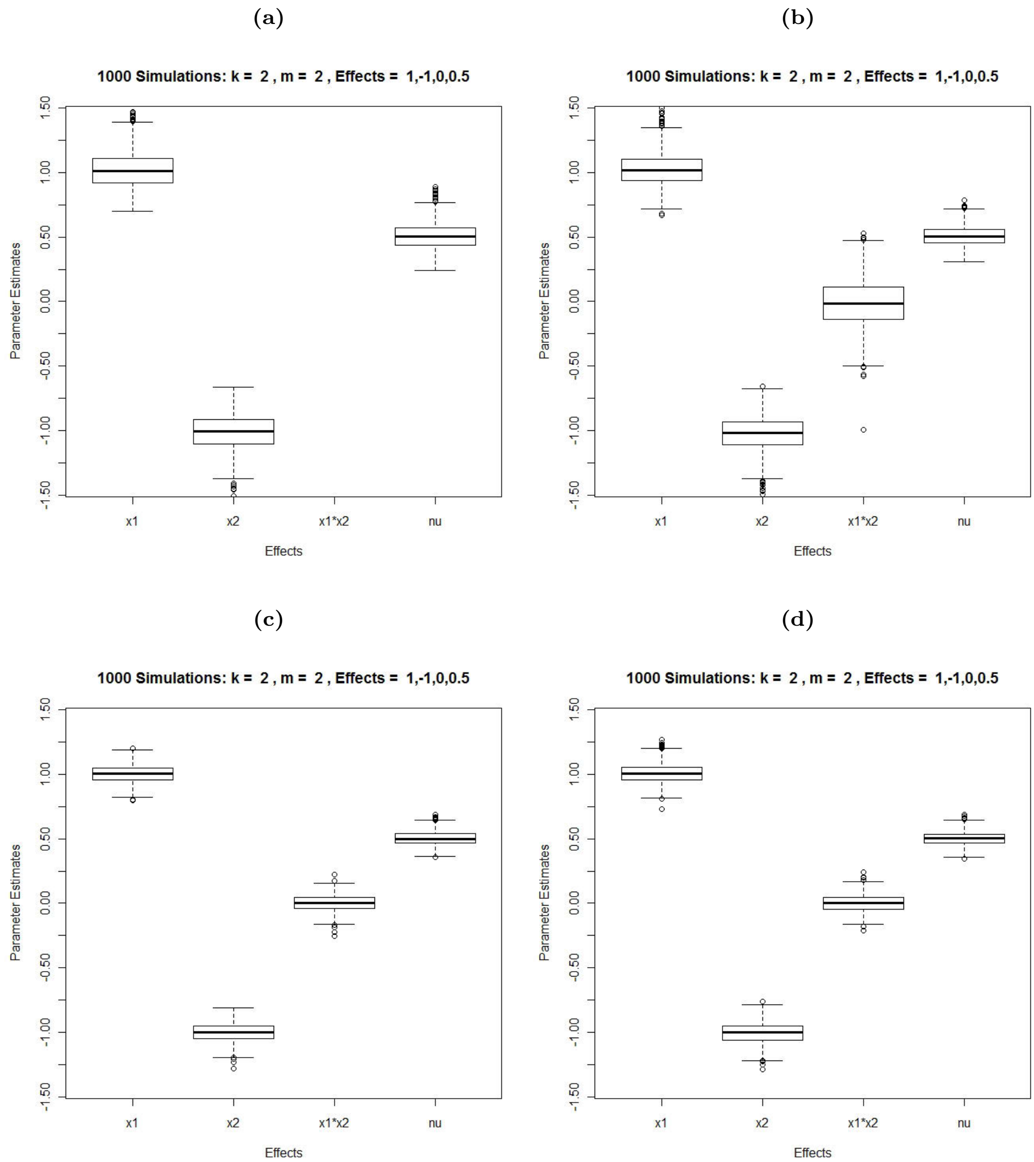

Figure 2.5: Simulation: estimating main effects and ν, design in (a) Table 2.5, (b) Table 2.6, (c) Table 2.11, and (d) Table 2.12.

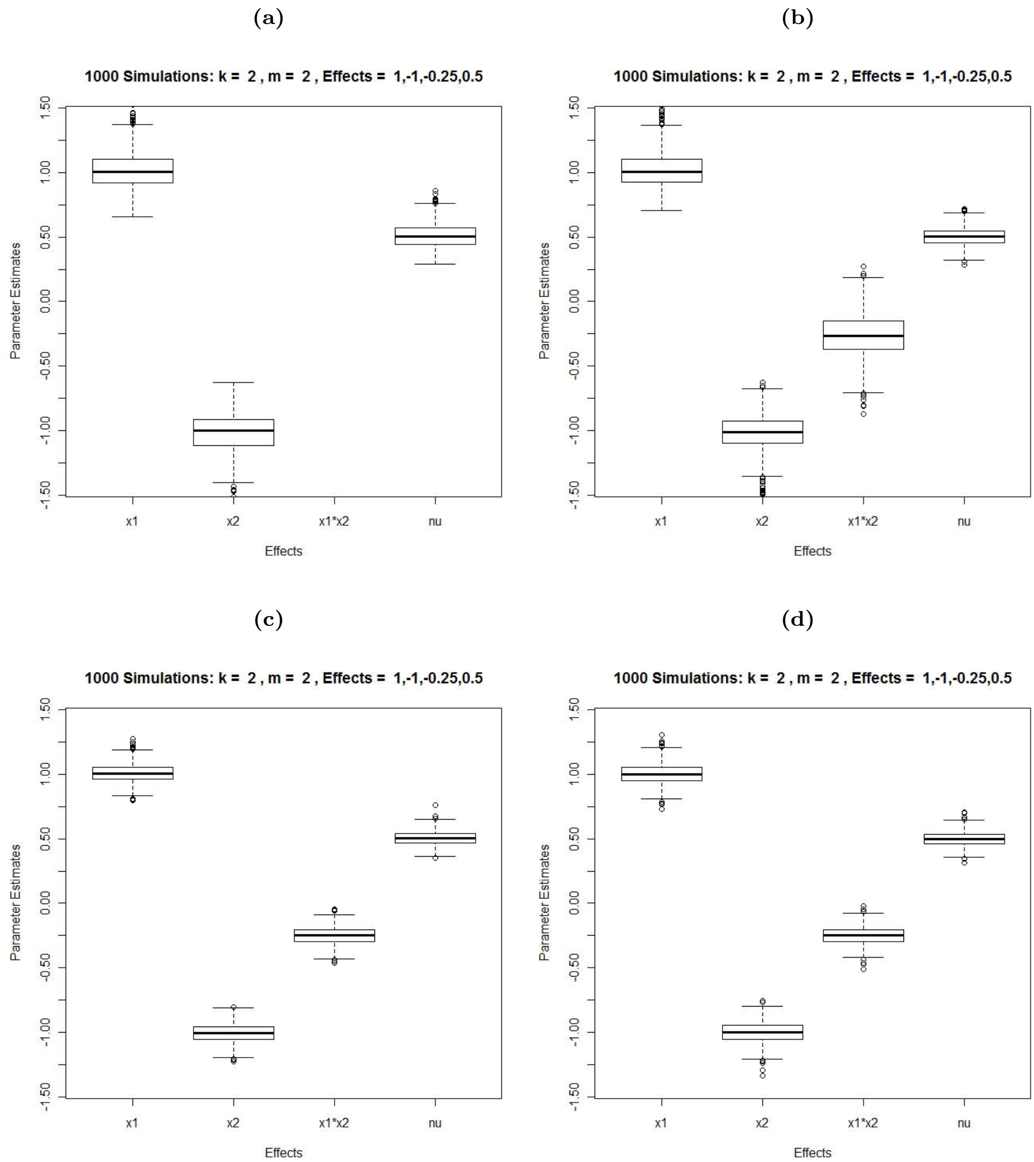

Figure 2.6: Simulation: estimating main effects, two–factor interactions, and ν, design in (a) Table 2.5, (b) Table 2.6, (c) Table 2.11, and (d) Table 2.12.

Parameter	Simulated Mean (Standard Error)	Theoretical Variance	Simulated Variance	Simulated Skewness (Standard Error)
Design in Table 2.5				
τ_1	1.02128(0.00449)	0.00718	0.02017	0.67056(0.07734)
τ_2	−1.01634(0.00448)	0.00718	0.02005	−0.54560(0.07734)
τ_{12}	Not Estimable			
ν	0.50828(0.00314)	0.00565	0.00984	0.38768(0.07734)
Design in Table 2.6				
τ_1	1.02945(0.00439)	0.00355	0.01926	0.68980(0.07734)
τ_2	−1.02916(0.00435)	0.00355	0.01895	−0.70955(0.07734)
τ_{12}	−0.01589(0.00589)	0.00638	0.03464	−0.18825(0.07734)
ν	0.50870(0.00242)	0.00185	0.00586	0.27522(0.07734)
Design in Table 2.11				
τ_1	1.00271(0.00217)	0.00253	0.00471	0.05708(0.07734)
τ_2	−1.00320(0.00221)	0.00199	0.00487	−0.15041(0.07734)
τ_{12}	−0.00035(0.00198)	0.00173	0.00392	−0.13595(0.07734)
ν	0.50460(0.00169)	0.00181	0.00285	0.23190(0.07734)
Design in Table 2.12				
τ_1	1.00910(0.00236)	0.00256	0.00557	0.18686(0.07734)
τ_2	−1.00539(0.00246)	0.00196	0.00608	−0.19718(0.07734)
τ_{12}	0.00025(0.00199)	0.00145	0.00397	0.06125(0.07734)
ν	0.50319(0.00167)	0.00178	0.00280	0.19896(0.07734)

Table 2.13: Summary statistics for $\tau_1 = 1, \tau_2 = -1, \tau_{12} = 0$, and $\nu = 0.5$.

Parameter	Simulated Mean (Standard Error)	Theoretical Variance	Simulated Variance	Simulated Skewness (Standard Error)
Design in Table 2.5				
τ_1	1.01915(0.00458)	0.00718	0.02102	0.56405(0.07734)
τ_2	−1.01681(0.00460)	0.00718	0.02112	−0.49041(0.07734)
τ_{12}	Not Estimable			
ν	0.50874(0.00297)	0.00565	0.00882	0.39822(0.07734)
Design in Table 2.6				
τ_1	1.02073(0.00448)	0.00355	0.02008	0.68262(0.07734)
τ_2	−1.02401(0.00437)	0.00355	0.01913	−0.49267(0.07734)
τ_{12}	−0.26706(0.00544)	0.00615	0.02964	−0.18378(0.07734)
ν	0.50377(0.00222)	0.00186	0.00494	0.21732(0.07734)
Design in Table 2.11				
τ_1	1.00766(0.00237)	0.00266	0.00560	0.19277(0.07734)
τ_2	−1.00705(0.00227)	0.00199	0.00514	−0.16026(0.07734)
τ_{12}	−0.25287(0.00205)	0.00179	0.00419	−0.02871(0.07734)
ν	0.50522(0.00172)	0.00181	0.00295	0.17568(0.07734)
Design in Table 2.12				
τ_1	1.00327(0.00250)	0.00272	0.00624	0.20013(0.07734)
τ_2	−1.00390(0.00252)	0.00196	0.00633	−0.21386(0.07734)
τ_{12}	−0.25104(0.00211)	0.00150	0.00445	−0.02041(0.07734)
ν	0.50264(0.00175)	0.00178	0.00305	0.27551(0.07734)

Table 2.14: Summary statistics for $\tau_1 = 1, \tau_2 = -1, \tau_{12} = -0.25$, and $\nu = 0.5$.

2.A Proof that $\mathbf{j_L^T z} = 0$ for the Davidson Ties Model

We begin by recalling that

$$z_i = \sum_{i_2 \neq i} w_{ii_2} + \frac{1}{2} w_{0|ii_2} - \frac{sn_{ii_2}\widehat{\pi}_i(1 + \frac{\widehat{\pi_{i_2}\nu}}{2\sqrt{\widehat{\pi}_i\widehat{\pi_{i_2}}}})}{\widehat{\pi}_i + \widehat{\pi_{i_2}} + \widehat{\nu}\sqrt{\widehat{\pi}_i\widehat{\pi_{i_2}}}}.$$

Now, the vector $\boldsymbol{z}$ contains the values for z_i for each possible item T_i. Then

$$\begin{aligned}
\boldsymbol{j}_L^T \boldsymbol{z} &= \sum_{i=1}^{t} z_i \\
&= \sum_{i_1 \neq i_2} w_{i|ii_2} + \frac{1}{2} w_{0(i,i_2)|ii_2} - \frac{sn_{ii_2}\widehat{\pi}_i(1 + \frac{\widehat{\pi_{i_2}\nu}}{2\sqrt{\widehat{\pi}_i\widehat{\pi_{i_2}}}})}{\widehat{\pi}_i + \widehat{\pi_{i_2}} + \widehat{\nu}\sqrt{\widehat{\pi}_i\widehat{\pi_{i_2}}}} \\
&= \sum_{C} (w_{i|ii_2} + w_{i_2|ii_2} + w_{0(i,i_2)|ii_2}) - \sum_{C} \frac{sn_C(\widehat{\pi_{i_1}} + \widehat{\pi_{i_2}} + \frac{\widehat{\pi_{i_1}}\widehat{\pi_{i_2}}\widehat{\nu}}{\sqrt{\widehat{\pi_{i_1}}\widehat{\pi_{i_2}}}})}{\widehat{\pi_{i_1}} + \widehat{\pi_{i_2}} + \widehat{\nu}\sqrt{\widehat{\pi_{i_1}}\widehat{\pi_{i_2}}}} \\
&= \sum_{C} sn_C - \sum_{C} \frac{sn_C(\widehat{\pi_{i_1}} + \widehat{\pi_{i_2}} + \widehat{\nu}\sqrt{\widehat{\pi_{i_1}}\widehat{\pi_{i_2}}})}{\widehat{\pi_{i_1}} + \widehat{\pi_{i_2}} + \widehat{\nu}\sqrt{\widehat{\pi_{i_1}}\widehat{\pi_{i_2}}}} \\
&= \sum_{C} sn_C - \sum_{C} sn_C \\
&= 0,
\end{aligned}$$

as required.

2.B Proof that the Davidson Ties Model does not violate El-Helbawy and Bradley [1978] Conditions

In order to apply the results relating to associated populations, we need to show that $C(\boldsymbol{\pi}_0, \nu)$ is positive definite, as El-Helbawy and Bradley [1978] did.

■ **THEOREM 2.B.1.**
The C matrix for the estimation of a set of contrasts $B_h\boldsymbol{\gamma}$ and ν, where

$$B = \begin{bmatrix} B_h & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix}$$

is positive definite.

Proof. El-Helbawy and Bradley [1978] show that $C(\boldsymbol{\pi}_0)_{\text{B–T}}$ is positive definite. Then the eigenvalues of $C(\boldsymbol{\pi}_0)_{\text{B–T}}$, $\lambda_1, \ldots, \lambda_p$ are all positive. If we consider the matrix

$$C(\boldsymbol{\pi}_0, \nu)_{\text{DAV}} = \begin{bmatrix} \frac{2}{2+\nu} C(\boldsymbol{\pi}_0)_{\text{B–T}} & \mathbf{0} \\ \mathbf{0} & \frac{2}{\nu(2+\nu)^2} \end{bmatrix},$$

Then

$$\det(C(\boldsymbol{\pi}_0, \nu)_{\text{DAV}} - I_{p+1}) = \frac{2}{2+\nu} \det(C(\boldsymbol{\pi}_0)_{\text{B–T}} - I_p) \times (\frac{2}{\nu(2+\nu)^2} - 1),$$

will have roots at $\lambda_1, \ldots, \lambda_p$, $\frac{2}{\nu(2+\nu)^2}$. We already know that $\lambda_1, \ldots, \lambda_p$ are positive, and $\frac{2}{\nu(2+\nu)^2}$ is positive for all $\nu > 0$. Since we assume that $\nu > 0$ anyway, all of the eigenvalues are positive. Therefore the C matrix for the Davidson ties model is positive definite. □

Chapter 3

The Generalised Davidson Ties Model

In Chapter 2, we found optimal designs for experiments when the Davidson ties model is used for choice sets of size 2. In Section 1.1, we introduced the MNL model as a generalisation of the Bradley–Terry model, which allows for an arbitrary number of options in a choice set.

This chapter introduces a generalisation to the MNL model to accommodate ties. This generalisation is analogous to the use of the Davidson ties model as a generalisation of the Bradley–Terry model to accommodate ties. We will first set up the model, the probability and likelihood functions and derive the information matrix for the estimation of a set of contrasts and the ties parameter.

Once we have established the properties of this model, we will then use the information matrix to show that the designs that are optimal when the MNL model is used are also optimal when the generalised Davidson ties model is used. Finally, we will look at some simulations of the generalised Davidson ties model.

3.1 Estimation of the generalised Davidson ties model

We begin this section by returning to the experiment in Example 2.0.12, looking at how we can allow for ties when we have larger choice sets.

■ **EXAMPLE 3.1.1.**
Consider the experiment introduced in Example 2.0.12. Suppose that we create choice sets with three items each. One such choice set is $\{00, 01, 10\}$. Then if we allow respondents to state that a subset of these items are equally attractive then there are 7 different outcomes arising from the choice set

- To state a preference for a single item, 00, 01 or 10,
- To state that a pair of items are equally attractive, $\{00, 01\}$, $\{00, 10\}$ or $\{01, 10\}$, and
- To state that all three items are equally attractive, $\{00, 01, 10\}$. □

Here we notice that the respondent is not only permitted to find pairs of items in the choice set equally attractive, but is also permitted to state that larger subsets of the items in the choice set are equally attractive. In his paper, Davidson argues that the merit of finding a set of items equally attractive is proportional to the geometric mean of the item merits. We will assume, as Davidson did, that the proportionality is constant across choice sets, and is strictly positive. If this constant is equal to zero then this means that no respondent has stated that any of the items in any of the choice sets are equally preferable to another item in the choice set, and the MNL model should be used instead.

■ **EXAMPLE 3.1.2.**

Consider the experiment introduced in Example 3.1.1. If we assign merits π_{00}, π_{01}, and π_{10} to the items in the choice set then we obtain merits $\nu\sqrt{\pi_{00}\pi_{01}}$, $\nu\sqrt{\pi_{00}\pi_{10}}$ and $\nu\sqrt{\pi_{01}\pi_{10}}$ for finding pairs of items equally preferable, and the merit of finding all three items in the choice set equally attractive $\nu\sqrt[3]{\pi_{00}\pi_{01}\pi_{10}}$. Let

$$D_{\{00,01,10\}} = \pi_{00} + \pi_{01} + \pi_{10} + \nu\sqrt{\pi_{00}\pi_{01}} + \nu\sqrt{\pi_{00}\pi_{10}} + \nu\sqrt{\pi_{01}\pi_{10}} + \nu\sqrt[3]{\pi_{00}\pi_{01}\pi_{10}}.$$

Then the selection probabilities of each of the types of choices take the form

$$\begin{aligned} P(00|\{00,01,10\}) &= \frac{\pi_{00}}{D_{\{00,01,10\}}}, \\ P(\{00,01\}|\{00,01,10\}) &= \frac{\nu\sqrt{\pi_{00}\pi_{01}}}{D_{\{00,01,10\}}}, \end{aligned}$$

and

$$P(\{00,01,10\}|\{00,01,10\}) = \frac{\nu\sqrt[3]{\pi_{00}\pi_{01}\pi_{10}}}{D_{\{00,01,10\}}}.$$

If we set $\nu = 0$ then we are saying that respondents will always be able to choose a single item as best. These probabilities then become

$$\begin{aligned} P(00|\{00,01,10\}) &= \frac{\pi_{00}}{\pi_{00} + \pi_{01} + \pi_{10}}, \\ P(\{00,01\}|\{00,01,10\}) &= 0, \end{aligned}$$

and

$$P(\{00,01,10\}|\{00,01,10\}) = 0,$$

that is, we are left with the MNL model. □

In general, let the merit of choosing item T_i be π_i, the merit of finding the items T_{i_1} and T_{i_2} equally attractive be $\nu\sqrt{\pi_{i_1}\pi_{i_2}}$. Let the merit of finding the items T_{i_1}, T_{i_2} and T_{i_3} equally attractive be $\nu\sqrt[3]{\pi_{i_1}\pi_{i_2}\pi_{i_3}}$. We continue this until we get to the respondent finding all of the m items in the choice set equally attractive, which will have merit $\nu\sqrt[m]{\pi_{i_1}\pi_{i_2}\ldots\pi_{i_m}}$. As in the MNL model, to obtain the probability of making a particular decision, we divide the merit of that decision by the sum of the merits of all possible decisions from the choice set. For a choice set $C = \{T_{i_1}, T_{i_2}, \ldots, T_{i_m}\}$, we denote the sum of the merits for each of the possible decisions by D_C. That is,

$$D_C = \sum_{a=1}^{m} \pi_{i_a} + \sum_{x=2}^{m} \sum_{\{T_{j_1},\ldots,T_{j_x}\}\subseteq C} \nu\sqrt[x]{\pi_{j_1}\ldots\pi_{j_x}}.$$

We can then express the probabilities for each decision as

$$\begin{aligned}
P(T_{i_1}|C) &= \frac{\pi_{i_1}}{D_C}, \\
P(\{T_{i_1}, T_{i_2}\}|C) &= \frac{\nu\sqrt{\pi_{i_1}\pi_{i_2}}}{D_C}, \\
P(\{T_{i_1}, T_{i_2}, T_{i_3}\}|C) &= \frac{\nu\sqrt[3]{\pi_{i_1}\pi_{i_2}\pi_{i_3}}}{D_C}, \\
&\vdots \\
P(\{T_{i_1}, T_{i_2}, \dots, T_{i_m}\}|C) &= \frac{\nu\sqrt[m]{\pi_{i_1}\pi_{i_2}\dots\pi_{i_m}}}{D_C}.
\end{aligned}$$

As before, we can define indicator variables $\boldsymbol{w}$ to represent whether a particular decision was made by a particular respondent, α, or not. We let

$$\begin{aligned}
w_{i_1|C,\alpha} &= \begin{cases} 1, & \text{if respondent } \alpha \text{ selected item } T_{i_1} \\ & \text{when presented with choice set } C, \\ 0, & \text{otherwise,} \end{cases} \\
w_{\{i_1,i_2\}|C,\alpha} &= \begin{cases} 1, & \text{if respondent } \alpha \text{ found items } T_{i_1} \text{ and } T_{i_2} \text{ equally} \\ & \text{attractive when presented with choice set } C, \\ 0, & \text{otherwise,} \end{cases} \\
&\vdots \\
w_{\{i_1,\dots,i_m\}|C,\alpha} &= \begin{cases} 1, & \text{if respondent } \alpha \text{ found all items in the} \\ & \text{choice set } C = \{T_{i_1}, \dots, T_{i_m}\} \text{ equally attractive,} \\ 0, & \text{otherwise,} \end{cases}
\end{aligned}$$

where for a given choice set and respondent α, we let only one of the ws be equal to 1, depending on the respondent's choice. This implies that there will be no repeated choice sets for any respondent, and no opt–out process. Then, for respondent α, the probability density function for the response to choice set $C = \{T_{i_1}, \dots, T_{i_m}\}$ is

$$f_{C,\alpha}(\boldsymbol{\pi}, \boldsymbol{w}, \nu) = \frac{1}{D_C^{n_C}} \times \prod_{i|T_i \in C} \pi_i^{w_{i|C,\alpha}} \times \prod_{x=2}^{m} \prod_{\{T_{j_1},\dots,T_{j_x}\}\subseteq C} (\nu\sqrt[x]{\pi_{j_1}\dots\pi_{j_x}})^{w_{\{j_1,\dots,j_x\}|C,\alpha}},$$

where n_C is an indicator variable which equals 1 if the choice set C is included in the experiment and 0 if it is not. For consistency, we will also let $w_{\{j_1,\dots,j_x\}|C,\alpha} = 0$, where $\{j_1, \dots, j_x\} \subseteq C$, if the choice set C does not appear in the experiment. The derivative of $\ln(f_{C,\alpha}(\boldsymbol{\pi}, \boldsymbol{w}, \nu))$ with respect to π_i is

$$\begin{aligned}
&\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{\pi}, \boldsymbol{w}, \nu))}{\partial \pi_i} \\
&= \frac{w_{i|C,\alpha}}{\pi_i} + \sum_{x=2}^{m} \sum_{\{T_{j_2},\dots,T_{j_x}\}\subseteq C} \frac{w_{\{i,j_2,\dots,j_x\}|C,\alpha}}{x\pi_i} - \frac{n_C}{D_C}\left(1 + \sum_{x=2}^{m} \sum_{\{T_i,T_{j_2},\dots,T_{j_x}\}\subseteq C} \frac{\nu\pi_{j_2}\dots\pi_{j_x}}{x\sqrt[x]{\pi_i\pi_{j_2}\dots\pi_{j_x}}}\right) \\
&= 0 \qquad \text{if } T_i \notin C,
\end{aligned}$$

and the derivative of $\ln(f_{C,\alpha}(\boldsymbol{\pi}, \boldsymbol{w}, \nu))$ with respect to ν is

$$\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{\pi}, \boldsymbol{w}, \nu))}{\partial \nu} = \sum_{x=2}^{m} \sum_{\{T_{j_1},\dots,T_{j_x}\}\subseteq C} \frac{w_{\{j_1,\dots,j_x\}|C,\alpha}}{\nu} - \frac{n_C}{D_C}\sum_{x=2}^{m} \sum_{\{T_{j_1},\dots,T_{j_x}\}\subseteq C} \sqrt[x]{\pi_{j_1}\dots\pi_{j_x}}.$$

We will use these derivatives later to derive an expression for the information matrix for this model. Before we derive the information matrix however, we will consider the MLEs for this model.

Since the likelihood function is the product of the density function for a respondent and choice set over all possible choice sets and over all of the respondents, we have

$$\begin{aligned} L(\boldsymbol{\pi}, \boldsymbol{w}, \nu) &= \prod_{\alpha=1}^{s} \prod_{C} f_{C,\alpha}(\boldsymbol{\pi}, \boldsymbol{w}, \nu) \\ &= \prod_{C} \frac{1}{D_C^{sn_C}} \left(\prod_{i|T_i \in C} \pi_i^{w_{i|C}} \times \sum_{x=2}^{m} \prod_{\{T_{j_1},\ldots,T_{j_x}\}\subseteq C} (\nu \sqrt[x]{\pi_{j_1} \ldots \pi_{j_x}})^{w_{\{j_1,\ldots,j_x\}|C}} \right). \end{aligned}$$

Notice that n_C is not subscripted by respondent. This is because we will assume that all respondents are presented with the same set of choice sets. Once again, we let

$$w_{\{j_1,\ldots,j_x\}|C} = \sum_{\alpha=1}^{s} w_{\{j_1,\ldots,j_x\}|C,\alpha}.$$

To maximise the likelihood function subject to the constraints of the model, we need to set up a Lagrangian function to incorporate the constraints. For purposes of convergence, we enforce the normalising constraint typically present in the MNL model

$$\sum_{i=1}^{t} \ln(\pi_i) = 0.$$

We also constrain the contrasts that we assume to be negligible to be equal to 0. Suppose that the matrix B_a contains the coefficients of these contrasts. Then

$$B_a \boldsymbol{\gamma} = 0,$$

where once again $\boldsymbol{\gamma}$ is a vector containing $\gamma_i = \ln(\pi_i)$ for $i = 1, 2, \ldots, t$. This gives the Lagrangian

$$\begin{aligned} G(\boldsymbol{\pi}, \boldsymbol{w}, \nu) = \sum_{C} \Bigg(&\sum_{i|T_i \in C} w_{i|C,\alpha} \ln(\pi_i) + \sum_{x=2}^{m} \sum_{\{T_{j_1},\ldots,T_{j_x}\}\subseteq C} w_{\{j_1,\ldots,j_x\}|C,\alpha} \left(\ln(\nu) + \frac{1}{x} \sum_{q=1}^{x} \ln(\pi_{j_q}) \right) \\ &- sn_C \ln D_C \Bigg) + \kappa_1 \sum_{i=1}^{t} \ln(\pi_i) + [\kappa_2, \ldots, \kappa_{a+1}] B_a \ln(\boldsymbol{\pi}), \end{aligned}$$

where $\kappa_1, \ldots, \kappa_{a+1}$ are Lagrange multipliers. If we differentiate $G(\boldsymbol{\pi}, \boldsymbol{w}, \nu)$ with respect to π_i, we obtain

$$\begin{aligned} \frac{\partial G(\boldsymbol{\pi}, \boldsymbol{w}, \nu)}{\partial \pi_i} &= \sum_{i|T_i \in C} \left(\frac{w_{i|C}}{\pi_i} + \sum_{x=2}^{m} \sum_{\{T_i, T_{j_2},\ldots,T_{j_x}\}\subseteq C} \frac{w_{\{i,j_2,\ldots,j_x\}|C}}{x\pi_i} - sn_C \frac{\partial D_C}{\partial \pi_i} \right) \\ &\quad + \frac{\kappa_1}{\pi_i} + \frac{1}{\pi_i} \sum_{u=1}^{a} \kappa_{u+1} (B_a)_{ui}. \end{aligned}$$

If we differentiate $G(\boldsymbol{\pi}, \boldsymbol{w}, \nu)$ with respect to ν, we obtain

$$\frac{\partial G(\boldsymbol{\pi}, \boldsymbol{w}, \nu)}{\partial \nu} = \sum_{C} \left(\sum_{x=2}^{m} \sum_{\{T_{j_1},\ldots,T_{j_x}\}\subseteq C} \frac{w_{\{j_1,\ldots,j_x\}|C}}{\nu} - sn_C \frac{\partial D_C}{\partial \nu} \right).$$

As usual, we obtain the MLEs by setting these equations equal to 0 and solving simultaneously.

This problem can be simplified by using matrix notation. Suppose that we let

$$z_i = \sum_{i|T_i \in C} \left(w_{i|C} + \sum_{x=2}^{m} \sum_{\{T_i, T_{j_2}, \ldots, T_{j_x}\} \subseteq C} \frac{1}{x} w_{\{i, j_2, \ldots, j_x\}|C} - sn_C \widehat{\pi}_i \frac{\partial D_C}{\partial \pi_i} \right).$$

Then by multiplying $\frac{\partial G(\boldsymbol{\pi}, \boldsymbol{w}, \nu)}{\partial \pi_i}$ by each π_i in turn, we get

$$z_i + \kappa_1 + \sum_{u=1}^{a} \kappa_{u+1} (B_a)_{ui} = 0, \qquad \text{for } i = 1, 2, \ldots, t.$$

This gives the system of equations

$$\boldsymbol{z} + \kappa_1 \boldsymbol{j}_L + B_a^T \boldsymbol{\kappa} = \mathbf{0}, \tag{3.1}$$

where $\boldsymbol{z} = (z_1, z_2, \ldots, z_t)^T$ and $\boldsymbol{\kappa} = (\kappa_2, \kappa_3, \ldots, \kappa_{a+1})^T$. Similarly, if we let

$$p = \sum_{C} \left(\sum_{x=2}^{m} \sum_{\{T_{j_1}, \ldots, T_{j_x}\} \subseteq C} w_{\{j_1, \ldots, j_x\}|C} - \nu sn_C \frac{\partial D_C}{\partial \nu} \right),$$

then we obtain $p = 0$ as the other equation to solve.

If we pre–multiply Equation 3.1 by $\boldsymbol{j}_L^T$ we obtain $\kappa_1 = 0$, since $\boldsymbol{j}_L^T \boldsymbol{z}$ is shown to be equal to 0 in Appendix 3.A, and the rows of B_a are the coefficients of contrasts so $\boldsymbol{j}_L^T B_a^T = (B_a \boldsymbol{j}_L)^T = \mathbf{0}$. Pre–multiplying Equation 3.1 by B_a, we obtain

$$\begin{aligned} B_a \boldsymbol{z} + \kappa_1 B_a \boldsymbol{j}_L + B_a B_a^T \boldsymbol{\kappa} &= \mathbf{0} \\ B_a \boldsymbol{z} + \boldsymbol{\kappa} &= \mathbf{0} \\ \boldsymbol{\kappa} &= -B_a \boldsymbol{z}. \end{aligned}$$

Substituting this into Equation 3.1, we get

$$(I - B_a^T B_a) \boldsymbol{z} = \mathbf{0},$$

and

$$p = 0$$

as the normal equations. We obtain the MLEs by solving these equations iteratively.

■ **EXAMPLE 3.1.3.**

Recall the experiment introduced in Example 3.1.1. Suppose that we present four choice sets to each respondent, $\{00, 01, 10\}$, $\{01, 00, 11\}$, $\{10, 11, 00\}$ and $\{11, 10, 01\}$ and that we are interested in the estimation of main effects and ν. Suppose that we present these choice sets to 150 respondents and obtain the summarised responses in Table 3.1. Then the Lagrangian for this

Option 1		Option 2		Option 3		T_1	T_2	T_3	$\{T_1, T_2\}$	$\{T_1, T_3\}$	$\{T_2, T_3\}$	All
0	0	0	1	1	0	5	19	54	10	21	20	21
0	1	0	0	1	1	4	4	73	5	25	22	17
1	0	1	1	0	0	26	61	0	27	9	15	12
1	1	1	0	0	1	47	23	8	28	15	10	19

Table 3.1: Responses for the experiment in Example 3.1.3.

estimation of the generalised Davidson ties model for this experiment is

$$
\begin{aligned}
&G(\boldsymbol{w}, \boldsymbol{\pi}, \nu) = \\
&5\ln(\pi_{00}) + 19\ln(\pi_{01}) + 54\ln(\pi_{10}) + (10+21+20+21)\ln(\nu) + \frac{10}{2}(\ln(\pi_{00}) + \ln(\pi_{01})) \\
&+\frac{21}{2}(\ln(\pi_{00})+\ln(\pi_{10}))+\frac{20}{2}(\ln(\pi_{01})+\ln(\pi_{10}))+\frac{21}{3}(\ln(\pi_{00})+\ln(\pi_{01})+\ln(\pi_{10}))-150\ln(D_{00,01,10}) \\
&+4\ln(\pi_{01}) + 4\ln(\pi_{00}) + 73\ln(\pi_{11}) + (5+25+22+17)\ln(\nu) + \frac{5}{2}(\ln(\pi_{01}) + \ln(\pi_{00})) \\
&+\frac{25}{2}(\ln(\pi_{01})+\ln(\pi_{11}))+\frac{22}{2}(\ln(\pi_{00})+\ln(\pi_{11}))+\frac{17}{3}(\ln(\pi_{01})+\ln(\pi_{00})+\ln(\pi_{11}))-150\ln(D_{01,00,11}) \\
&+26\ln(\pi_{10}) + 61\ln(\pi_{11}) + 0\ln(\pi_{00}) + (27+9+15+12)\ln(\nu) + \frac{27}{2}(\ln(\pi_{00}) + \ln(\pi_{01})) \\
&+\frac{9}{2}(\ln(\pi_{00})+\ln(\pi_{01}))+\frac{15}{2}(\ln(\pi_{00})+\ln(\pi_{01}))+\frac{12}{3}(\ln(\pi_{00})+\ln(\pi_{01})+\ln(\pi_{10}))-150\ln(D_{10,11,00}) \\
&+47\ln(\pi_{11}) + 23\ln(\pi_{10}) + 8\ln(\pi_{01}) + (28+15+10+19)\ln(\nu) + \frac{28}{2}(\ln(\pi_{00}) + \ln(\pi_{01})) \\
&+\frac{15}{2}(\ln(\pi_{00})+\ln(\pi_{01}))+\frac{10}{2}(\ln(\pi_{00})+\ln(\pi_{01})+\frac{19}{3}(\ln(\pi_{00})+\ln(\pi_{01})+\ln(\pi_{10}))-150\ln(D_{11,10,01}) \\
&+\kappa_1(\ln(\pi_{00}) + \ln(\pi_{01}) + \ln(\pi_{10}) + \ln(\pi_{11})) + \kappa_2(\ln(\pi_{00}) - \ln(\pi_{01}) - \ln(\pi_{10}) + \ln(\pi_{11})).
\end{aligned}
$$

We differentiate the Lagrangian with respect to each π_i in turn. For instance,

$$
\begin{aligned}
\frac{\partial G(\boldsymbol{w}, \boldsymbol{\pi}, \nu)}{\partial \pi_{00}} = \frac{200}{3\pi_{00}} + \frac{\kappa_1}{\pi_{00}} + \frac{\kappa_2}{\pi_{00}} - 150 \Bigg(&\frac{1 + \frac{\nu\pi_{01}}{\sqrt{\pi_{00}\pi_{01}}} + \frac{\nu\pi_{10}}{\sqrt{\pi_{00}\pi_{10}}} + \frac{\nu\pi_{01}\pi_{10}}{\sqrt[3]{\pi_{00}\pi_{01}\pi_{10}}}}{D_{00,01,10}} \\
&- \frac{1 + \frac{\nu\pi_{01}}{\sqrt{\pi_{00}\pi_{01}}} + \frac{\nu\pi_{11}}{\sqrt{\pi_{00}\pi_{11}}} + \frac{\nu\pi_{01}\pi_{11}}{\sqrt[3]{\pi_{00}\pi_{01}\pi_{11}}}}{D_{01,00,11}} - \frac{1 + \frac{\nu\pi_{10}}{\sqrt{\pi_{00}\pi_{10}}} + \frac{\nu\pi_{11}}{\sqrt{\pi_{00}\pi_{11}}} + \frac{\nu\pi_{10}\pi_{11}}{\sqrt[3]{\pi_{00}\pi_{10}\pi_{11}}}}{D_{10,11,00}} \Bigg),
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial G(\boldsymbol{w}, \boldsymbol{\pi}, \nu)}{\partial \nu} &= \frac{263}{\nu} - \frac{150(\sqrt{\pi_{00}\pi_{01}} + \sqrt{\pi_{00}\pi_{10}} + \sqrt{\pi_{01}\pi_{10}} + \sqrt[3]{\pi_{00}\pi_{01}\pi_{10}})}{D_{00,01,10}} \\
&- \frac{150(\sqrt{\pi_{00}\pi_{01}} + \sqrt{\pi_{00}\pi_{11}} + \sqrt{\pi_{01}\pi_{11}} + \sqrt[3]{\pi_{00}\pi_{01}\pi_{11}})}{D_{01,00,11}} \\
&- \frac{150(\sqrt{\pi_{00}\pi_{10}} + \sqrt{\pi_{00}\pi_{11}} + \sqrt{\pi_{10}\pi_{11}} + \sqrt[3]{\pi_{00}\pi_{10}\pi_{11}})}{D_{10,11,00}} \\
&- \frac{150(\sqrt{\pi_{01}\pi_{10}} + \sqrt{\pi_{01}\pi_{11}} + \sqrt{\pi_{10}\pi_{11}} + \sqrt[3]{\pi_{01}\pi_{10}\pi_{11}})}{D_{11,10,01}}.
\end{aligned}
$$

If we set each of these equal to 0 and solve iteratively then we obtain the MLEs for $\boldsymbol{\pi}$ and ν. If we let τ_1 be the main effect of the first attribute and τ_2 the main effect of the second attribute then we get

$$\hat{\tau_1} = 1.01 \qquad \hat{\tau_2} = 0.39 \qquad \hat{\nu} = 0.93. \qquad \square$$

3.2 Properties of the generalised Davidson ties model

In this section, we complete the construction of the information matrix for the estimation of the entries in $\boldsymbol{\pi}$ and ν. We begin by deriving expressions for the variances and covariances of the selection indicators, $\boldsymbol{w}$, introduced previously. We then use these expressions to simplify the information matrix.

Recall that the entries in $\boldsymbol{w}$ are selection indicators for the choice made by a respondent when presented with the choice set $C = \{T_{i_1}, T_{i_2}, \ldots, T_{i_m}\}$. These w_i have a Bernoulli distribution with expectations

$$\begin{aligned}
\mathcal{E}_\pi(w_{i_1|C,\alpha}) &= \frac{\pi_{i_1}}{D_C}, \\
\mathcal{E}_\pi(w_{\{i_1,i_2\}|C,\alpha}) &= \frac{\nu\sqrt{\pi_{i_1}\pi_{i_2}}}{D_C}, \\
\mathcal{E}_\pi(w_{\{i_1,i_2,i_3\}|C,\alpha}) &= \frac{\nu\sqrt[3]{\pi_{i_1}\pi_{i_2}\pi_{i_3}}}{D_C}, \\
&\vdots \\
\mathcal{E}_\pi(w_{\{i_1,i_2,\ldots,i_{m-1}\}|C,\alpha}) &= \frac{\nu\sqrt[m-1]{\pi_{i_1}\pi_{i_2}\cdots\pi_{i_{m-1}}}}{D_C},
\end{aligned}$$

and

$$\mathcal{E}_\pi(w_{\{i_1,i_2,\ldots,i_m\}|C,\alpha}) = \frac{\nu\sqrt[m]{\pi_{i_1}\pi_{i_2}\cdots\pi_{i_m}}}{D_C}. \tag{3.2}$$

The variances of these selection indicators are given by

$$\begin{aligned}
\mathrm{Var}_\pi(w_{i_1|C,\alpha}) &= \frac{\pi_{i_1}}{D_C} \times \frac{D_C - \pi_{i_1}}{D_C}, \\
\mathrm{Var}_\pi(w_{\{i_1,i_2\}|C,\alpha}) &= \frac{\nu\sqrt{\pi_{i_1}\pi_{i_2}}}{D_C} \times \frac{D_C - \nu\sqrt{\pi_{i_1}\pi_{i_2}}}{D_C}, \\
\mathrm{Var}_\pi(w_{\{i_1,i_2,i_3\}|C,\alpha}) &= \frac{\nu\sqrt[3]{\pi_{i_1}\pi_{i_2}\pi_{i_3}}}{D_C} \times \frac{D_C - \nu\sqrt[3]{\pi_{i_1}\pi_{i_2}\pi_{i_3}}}{D_C}, \\
&\vdots \\
\mathrm{Var}_\pi(w_{\{i_1,\ldots,i_{m-1}\}|C,\alpha}) &= \frac{\nu\sqrt[m-1]{\pi_{i_1}\cdots\pi_{i_{m-1}}}}{D_C} \times \frac{D_C - \nu\sqrt[m-1]{\pi_{i_1}\cdots\pi_{i_{m-1}}}}{D_C},
\end{aligned}$$

and

$$\mathrm{Var}_\pi(w_{\{i_1,\ldots,i_m\}|C,\alpha}) = \frac{\nu\sqrt[m]{\pi_{i_1}\cdots\pi_{i_m}}}{D_C} \times \frac{D_C - \nu\sqrt[m]{\pi_{i_1}\pi_{i_2}\cdots\pi_{i_m}}}{D_C}. \tag{3.3}$$

Next we derive covariances for the selection indicators. We assume that the selections made in two distinct choice sets are uncorrelated, and thus the selection indicators between choice sets have zero correlation.

We begin with the correlation between two selection indicators that both represent selections of a single item. We notice that it would not be possible to select both T_{i_1} alone and T_{i_2} alone, and therefore $\mathcal{E}_\pi(w_{i_1|C,\alpha}w_{i_2|C,\alpha}) = 0$. This yields

$$\begin{aligned}
\mathrm{Cov}_\pi(w_{i_1|C,\alpha}, w_{i_2|C,\alpha}) &= \mathcal{E}_\pi\Big(\big(w_{i_1|C,\alpha} - \mathcal{E}_\pi(w_{i_1|C,\alpha})\big)\big(w_{i_2|C,\alpha} - \mathcal{E}_\pi(w_{i_2|C,\alpha})\big)\Big) \\
&= 0 - \mathcal{E}_\pi(w_{i_1|C,\alpha})\mathcal{E}_\pi(w_{i_2|C,\alpha}) \\
&= \frac{-\pi_{i_1}\pi_{i_2}}{(D_{C,\alpha})^2}.
\end{aligned}$$

Then

$$\mathrm{Cov}_\pi(w_{i_1|C,\alpha}, w_{i_{1'}|C',\alpha}) = \begin{cases} \frac{-\pi_{i_1}\pi_{i_{1'}}}{(D_{C,\alpha})^2}, & \text{if } C = C' \text{ and } i_1 \neq i_{1'}, \\ \mathrm{Var}_\pi(w_{i_1|C,\alpha}), & \text{if } C = C' \text{ and } i_1 = i_{1'}, \\ 0, & \text{otherwise,} \end{cases}$$

where $C' = \{T_{i_{1'}}, T_{i_{2'}}, \ldots, T_{i_{m'}}\}$. We repeat this procedure for different selection indicators. This gives

$$\mathrm{Cov}_\pi(w_{j|C,\alpha}, w_{\{j_{1'},\ldots,j_{x'}\}|C',\alpha}) = \begin{cases} \frac{-\pi_j \nu \sqrt[x]{\pi_{j_{1'}} \times \ldots \times \pi_{i_{x'}}}}{(D_C)^2}, & \text{if } C = C', \\ 0, & \text{otherwise,} \end{cases}$$

and

$$\mathrm{Cov}_\pi(w_{\{j_1,\ldots,j_x\}|C,\alpha}, w_{\{j_{1'},\ldots,j_{u'}\}|C',\alpha}) = \begin{cases} \frac{-\nu^2 \sqrt[x]{\pi_{j_1} \times \ldots \times \pi_{i_x}} \sqrt[u]{\pi_{i_{1'}} \times \ldots \times \pi_{i_{u'}}}}{(D_C)^2}, & \text{if } C = C', \\ 0, & \text{otherwise,} \end{cases} \tag{3.4}$$

for $2 \leq x, y \leq m$.

Next we construct the information matrix for this model. As with the Davidson ties model, the construction will be easier if we partition the information matrix into four blocks. Thus let

$$I(\boldsymbol{\pi}, \nu) = \begin{pmatrix} I_{\pi\pi}(\boldsymbol{\pi}, \nu) & I_{\nu\pi}(\boldsymbol{\pi}, \nu) \\ I_{\pi\nu}(\boldsymbol{\pi}, \nu) & I_{\nu\nu}(\boldsymbol{\pi}, \nu) \end{pmatrix},$$

where $I_{\pi\pi}(\boldsymbol{\pi}, \nu)$ contains minus the expected value of second derivatives of the density function with respect to two of the entries in $\boldsymbol{\pi}$. $I_{\nu\nu}(\boldsymbol{\pi}, \nu)$ contains minus the expected value of the second derivative of the density function with respect to ν twice. $I_{\pi\nu}(\boldsymbol{\pi}, \nu)$ and $I_{\nu\pi}(\boldsymbol{\pi}, \nu)$ contains minus the expected value of the second derivatives with respect to one entry in $\boldsymbol{\pi}$ and ν, where $I_{\pi\nu}(\boldsymbol{\pi}, \nu) = I_{\nu\pi}(\boldsymbol{\pi}, \nu)^T$.

El-Helbawy and Bradley [1978] state that, under some mild regularity conditions, as given in Section 1.1, the $(i, j)^{\text{th}}$ entry of the information matrix for a discrete choice experiment without ties is

$$I(\boldsymbol{\pi})_{ij} = \sum_{q=1}^{t-1} \sum_{r=q+1}^{t} \frac{n_{qr}}{N} \mathcal{E}_\pi \left(\left(\frac{\partial \ln(f_{qr\alpha}(\boldsymbol{\pi}, \boldsymbol{w}))}{\partial \pi_i} \right) \left(\frac{\partial \ln(f_{qr\alpha}(\boldsymbol{\pi}, \boldsymbol{w}))}{\partial \pi_j} \right) \right).$$

We now use this expression, and the results in Equations 3.2, 3.3, and 3.4, to evaluate some generic cells in each block matrix. Since $m \geq 2$ in this case, we sum over the choice sets of size m rather than pairs of items, and modify the notation for n_{qr} and $f_{qr\alpha}(\boldsymbol{\pi}, \boldsymbol{w})$ accordingly. We will begin with $I_{\pi\pi}(\boldsymbol{\pi}, \nu)$. In this block matrix, we need to consider the off–diagonal and diagonal entries separately.

Let us begin with the off–diagonal entries of $I_{\pi\pi}(\boldsymbol{\pi}, \nu)$. Suppose that we consider a generic entry $I_{\pi\pi}(\boldsymbol{\pi}, \nu)_{ij}$, containing the product of the derivatives with respect to π_i and with respect to π_j. We begin with

$$I_{\pi\pi}(\boldsymbol{\pi}, \nu)_{ij} = \sum_C \frac{n_C}{N} \mathcal{E}_\pi \left(\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{\pi}, \boldsymbol{w}, \nu))}{\partial \pi_i} \frac{\partial \ln(f_{C,\alpha}(\boldsymbol{\pi}, \boldsymbol{w}, \nu))}{\partial \pi_j} \right).$$

The derivative of $f(\boldsymbol{\pi}, \boldsymbol{w}, \nu)$ with respect to π_i will be 0 unless T_i appears in the choice set C. Then we can restrict our summation to those choice sets that contain both T_i and T_j. This gives

$$I_{\pi\pi}(\boldsymbol{\pi}, \nu)_{ij} = \sum_{C|T_i,T_j \in C} \frac{n_C}{N} \mathcal{E}_\pi \left(\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{\pi}, \boldsymbol{w}, \nu))}{\partial \pi_i} \frac{\partial \ln(f_{C,\alpha}(\boldsymbol{\pi}, \boldsymbol{w}, \nu))}{\partial \pi_j} \right).$$

By observation, we obtain

$$\frac{\partial \ln(f_{C\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \nu))}{\partial \pi_i} = \frac{w_{i|C,\alpha}}{\pi_i} - \mathcal{E}_\pi\Big(\frac{w_{i|C,\alpha}}{\pi_i}\Big). \tag{3.5}$$

Then it follows that

$$\begin{aligned} I_{\pi\pi}(\boldsymbol{\pi}, \nu)_{ij} = \sum_{C|T_i,T_j\in C} \frac{n_C}{N}\mathrm{Cov}_\pi\Bigg(&\frac{w_{i|C,\alpha}}{\pi_i} + \sum_{x=2}^{m}\sum_{\{T_i,T_{j_2},\ldots,T_{j_x}\}\subseteq C} \frac{w_{\{i,j_2,\ldots,j_x\}|C,\alpha}}{x\pi_i}, \\ &\frac{w_{j|C,\alpha}}{\pi_j} + \sum_{u=2}^{m}\sum_{\{T_j,T_{j'_2},\ldots,T_{j'_u}\}\subseteq C} \frac{w_{\{j,j'_2,\ldots,j'_u\}|C,\alpha}}{u\pi_j}\Bigg). \end{aligned}$$

If we expand this covariance, we get

$$\begin{aligned} &I_{\pi\pi}(\boldsymbol{\pi}, \nu)_{ij} \\ &= \sum_{C|T_i,T_j\in C} \frac{n_C}{N\pi_i\pi_j}\Bigg(\mathrm{Cov}_\pi(w_{i|C,\alpha}, w_{j|C,\alpha}) + \sum_{u=2}^{m}\sum_{\{T_i,T_{j_2},\ldots,T_{j_u}\}\subseteq C}\frac{1}{u}\mathrm{Cov}_\pi(w_{i|C,\alpha}, w_{\{j,j'_2,\ldots,j'_u\}|C,\alpha}) \\ &\quad + \sum_{x=2}^{m}\sum_{\{T_i,T_{j_2},\ldots,T_{j_x}\}\subseteq C}\frac{1}{x}\mathrm{Cov}_\pi(w_{\{i,j_2,\ldots,j_x\}|C,\alpha}, w_{j|C,\alpha}) \\ &\quad + \sum_{x=2}^{m}\sum_{\{T_i,T_j,T_{j_3},\ldots,T_{j_x}\}\subseteq C}\frac{1}{x^2}\mathrm{Var}_\pi(w_{\{i,j,j_3,\ldots,j_x\}|C,\alpha}) \\ &\quad + \sum_{x=2}^{m}\sum_{u=2}^{m}\sum_{\substack{\{T_i,T_{j_2},\ldots,T_{j_x}\}\\ \neq\{T_j,T_{j'_2},\ldots,T_{j'_u}\}\subseteq C}}\frac{1}{xu}\mathrm{Cov}_\pi(w_{\{i,j_2,\ldots,j_x\}|C,\alpha}, w_{\{j,j'_2,\ldots,j'_u\}|C,\alpha})\Bigg). \end{aligned}$$

If we then substitute the Equations 3.3 and 3.4 and simplify, we obtain

$$\begin{aligned} I_{\pi\pi}(\boldsymbol{\pi}, \nu)_{ij} = \sum_{C|T_i,T_j\in C} \frac{n_C}{N\pi_i\pi_j D_C^2}\Bigg(& -\pi_i\pi_j - \sum_{u=2}^{m}\sum_{\{T_i,T_{j'_2},\ldots,T_{j'_u}\}\subseteq C}\frac{\pi_j\nu\sqrt[u]{\pi_i\pi_{j'_2}\ldots\pi_{j'_u}}}{u} \\ & - \sum_{x=2}^{m}\sum_{\{T_i,T_{j_2},\ldots,T_{j_x}\}\subseteq C}\frac{\pi_i\nu\sqrt[x]{\pi_j\pi_{j_2}\ldots\pi_{j_x}}}{x} \\ & + \sum_{x=2}^{m}\sum_{\{T_i,T_j,T_{j_3},\ldots,T_{j_x}\}}\frac{\nu\sqrt[x]{\pi_i\pi_j\pi_{j_3}\ldots\pi_{j_x}}(D_C - \nu\sqrt[x]{\pi_i\pi_j\pi_{j_3}\ldots\pi_{j_x}})}{x^2} \\ & - \sum_{x=2}^{m}\sum_{u=2}^{m}\sum_{\substack{\{T_i,T_{j_2},\ldots,T_{j_x}\}\\ \neq\{T_j,T_{j'_2},\ldots,T_{j'_u}\}}}\frac{\nu^2\sqrt[x]{\pi_i\pi_{j_2}\ldots\pi_{j_x}}\sqrt[u]{\pi_j\pi_{j'_2}\ldots\pi_{j'_u}}}{xy}\Bigg). \end{aligned}$$

Now let us consider a generic diagonal term in $I_{\pi\pi}(\boldsymbol{\pi}, \nu)$. Again, if we notice that the derivative of $\ln(f(\boldsymbol{\pi}, \boldsymbol{w}, \nu))$ with respect to π_i is 0 when T_i does not appear in the choice set, we obtain

$$I_{\pi\pi}(\boldsymbol{\pi}, \nu)_{ii} = \sum_{C|T_i\in C} \frac{n_C}{N}\mathcal{E}_\pi\Big(\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{\pi}, \boldsymbol{w}, \nu))}{\partial \pi_i}\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{\pi}, \boldsymbol{w}, \nu))}{\partial \pi_i}\Big).$$

Using Equation 3.5, this simplifies to

$$I_{\pi\pi}(\boldsymbol{\pi}, \nu)_{ii} = \sum_{C|T_i\in C} \frac{n_C}{N}\mathrm{Var}_\pi\Big(\frac{w_{i|C,\alpha}}{\pi_i} + \sum_{x=2}^{m}\sum_{\{T_i,T_{j_2},\ldots,T_{j_x}\}\subseteq C}\frac{w_{\{i,j_2,\ldots,j_x\}|C,\alpha}}{x\pi_i}\Big).$$

We can expand this variance to get

$$I_{\pi\pi}(\boldsymbol{\pi},\nu)_{ii}$$
$$=\sum_{C|T_i\in C}\frac{n_C}{N\pi_i^2}\left(\mathrm{Var}_\pi(w_{i|C,\alpha})+\sum_{u=2}^{m}\sum_{\{T_i,T_{j'_2},\ldots,T_{j'_u}\}\subseteq C}\frac{\mathrm{Cov}_\pi(w_{i|C,\alpha},w_{\{j,j'_2,\ldots,j'_u\}|C,\alpha})}{u}\right.$$
$$+\sum_{x=2}^{m}\sum_{\{T_i,T_{j_2},\ldots,T_{j_x}\}\subseteq C}\frac{\mathrm{Cov}_\pi(w_{\{i,j_2,\ldots,j_x\}|C,\alpha},w_{j|C,\alpha})}{x}+\sum_{x=2}^{m}\sum_{\{T_i,T_{j_2},\ldots,T_{j_x}\}\subseteq C}\frac{\mathrm{Var}_\pi(w_{\{i,j_2,\ldots,j_x\}|C,\alpha})}{x^2}$$
$$\left.+\sum_{x=2}^{m}\sum_{u=2}^{m}\sum_{\substack{\{T_i,T_{j_2},\ldots,T_{j_x}\}\\ \neq\{T_j,T_{j'_2},\ldots,T_{j'_u}\}\subseteq C}}\frac{\mathrm{Cov}_\pi(w_{\{i,j_2,\ldots,j_x\}|C,\alpha},w_{\{j,j'_2,\ldots,j'_u\}|C,\alpha})}{xu}\right).$$

Substituting Equations 3.3 and 3.4 and simplifying gives

$$I_{\pi\pi}(\boldsymbol{\pi},\nu)_{ii}$$
$$=\sum_{C|T_i\in C}\frac{n_C}{N\pi_i^2D_C^2}\left(\pi_i(D_C-\pi_i)+\sum_{x=2}^{m}\sum_{\{T_i,T_{j_2},\ldots,T_{j_x}\}\subseteq C}\frac{\nu\sqrt[x]{\pi_i\pi_{j_2}\ldots\pi_{j_x}}(D_C-\nu\sqrt[x]{\pi_i\pi_{j_2}\ldots\pi_{j_x}})}{x^2}\right.$$
$$-2\sum_{x=2}^{m}\sum_{\{T_i,T_{j_2},\ldots,T_{j_x}\}\subseteq C}\frac{\nu\pi_i\sqrt[x]{\pi_i\pi_{j_2}\ldots\pi_{j_x}}}{x}$$
$$\left.-\sum_{x=2}^{m}\sum_{u=2}^{m}\sum_{\substack{\{T_i,T_{j_2},\ldots,T_{j_x}\}\\ \neq\{T_i,T_{j'_2},\ldots,T_{j'_u}\}\subseteq C}}\frac{\nu^2\sqrt[x]{\pi_i\pi_{j_2}\ldots\pi_{j_x}}\sqrt[u]{\pi_i\pi_{j'_2}\ldots\pi_{j'_u}}}{xu}\right).$$

We can repeat this process for a generic entry of $I_{\pi\nu}(\boldsymbol{\pi},\nu)$, a $1\times t$ vector. In the same way as before we exclude the choice sets that do not include T_i, which will give a zero summand, and obtain

$$I_{\pi\nu}(\boldsymbol{\pi},\nu)_i = \sum_{C|T_i\in C}\frac{n_C}{N}\mathcal{E}_\pi\left(\frac{\partial\ln(f_{C,\alpha}(\boldsymbol{\pi},\boldsymbol{w}))}{\partial\pi_i}\frac{\partial\ln(f_{C,\alpha}(\boldsymbol{\pi},\boldsymbol{w}))}{\partial\nu}\right).$$

If we take into account Equation 3.5 and that

$$\frac{\partial\ln(f_{C\alpha}(\boldsymbol{\pi},\boldsymbol{w},\nu))}{\partial\nu}=\sum_{x=2}^{m}\sum_{\{T_{j_1},\ldots,T_{j_x}\}\subseteq C}\frac{w_{\{j_1,\ldots,j_x\}|C\alpha}}{\nu}-\mathcal{E}_\pi\left(\sum_{x=2}^{m}\sum_{\{T_{j_1},\ldots,T_{j_x}\}\subseteq C}\frac{w_{\{j_1,\ldots,j_x\}|C\alpha}}{\nu}\right),\tag{3.6}$$

then

$$I_{\pi\nu}(\boldsymbol{\pi},\nu)_i$$
$$=\sum_{C|T_i\in C}\frac{n_C}{N}\mathrm{Cov}_\pi\left(\frac{w_{i|C,\alpha}}{\pi_i}+\sum_{x=2}^{m}\sum_{\{T_i,T_{j_2},\ldots,T_{j_x}\}\subseteq C}\frac{w_{\{i,j_2,\ldots,j_x\}|C,\alpha}}{x\pi_i},\sum_{u=2}^{m}\sum_{\{T_{j'_1},\ldots,T_{j'_u}\}\subseteq C}\frac{w_{\{j'_1,\ldots,j'_u\}|C,\alpha}}{\nu}\right).$$

Again, we expand these covariances to obtain

$$\begin{aligned} I_{\pi\nu}(\boldsymbol{\pi},\nu)_i &= \sum_{C|T_i\in C}\frac{n_C}{N\pi_i\nu}\Bigg(\sum_{u=2}^{m}\sum_{\{T_i,T_{j'_2},\dots,T_{j'_u}\}\subseteq C}\mathrm{Cov}_\pi(w_{i|C,\alpha},w_{\{j'_1,\dots,j'_u\}|C,\alpha}) \\ &\quad+\sum_{x=2}^{m}\sum_{\{T_i,T_{j_2},\dots,T_{j_x}\}\subseteq C}\frac{1}{x}\mathrm{Var}_\pi(w_{\{i,j_2,\dots,j_x\}|C,\alpha}) \\ &\quad+\sum_{x=2}^{m}\sum_{u=2}^{m}\sum_{\substack{\{T_i,T_{j_2},\dots,T_{j_x}\}\\ \neq\{T_{j'_1},\dots,T_{j'_u}\}\subseteq C}}\frac{1}{x}\mathrm{Cov}_\pi(w_{\{i,j_2,\dots,j_x\}|C,\alpha},w_{\{j'_1,\dots,j'_u\}|C,\alpha})\Bigg). \end{aligned}$$

Substituting Equations 3.3 and 3.4 and simplifying gives

$$\begin{aligned} I_{\pi\nu}(\boldsymbol{\pi},\nu)_i &= \sum_{i|T_i\in C}\frac{n_C}{N\pi_i\nu D_C^2}\Bigg(-\sum_{u=2}^{m}\sum_{\{T_{j'_1},T_{j'_2},\dots,T_{j'_u}\}\subseteq C}\pi_i\nu\sqrt[u]{\pi_{j'_1}\pi_{j'_2}\dots\pi_{j'_u}} \\ &\quad+\sum_{x=2}^{m}\sum_{\{T_i,T_{j_2},\dots,T_{j_x}\}\subseteq C}\frac{\nu\sqrt[x]{\pi_i\pi_{j_2}\dots\pi_{j_x}}(D_C-\nu\sqrt[x]{\pi_i\pi_{j_2}\dots\pi_{j_x}})}{x} \\ &\quad-\sum_{x=2}^{m}\sum_{u=2}^{m}\sum_{\substack{\{T_i,T_{j_2},\dots,T_{j_x}\}\\ \neq\{T_{j'_1},T_{j'_2},\dots,T_{j'_u}\}\subseteq C}}\frac{\nu^2\sqrt[x]{\pi_i\pi_{j_2}\dots\pi_{j_x}}\sqrt[u]{\pi_{j'_1}\pi_{j'_2}\dots\pi_{j'_u}}}{x}\Bigg). \end{aligned}$$

Finally, we look at the single element $I_{\nu\nu}(\boldsymbol{\pi},\nu)$. We begin with

$$I_{\nu\nu}(\boldsymbol{\pi},\nu) = \sum_C\frac{n_C}{N}\mathcal{E}_\pi\left(\left(\frac{\partial\ln(f_{C,\alpha}(\boldsymbol{\pi},\boldsymbol{w}))}{\partial\nu}\right)^2\right).$$

In this case, there are no derivatives with respect to any π_i, so no choice sets are excluded from the summation. Thus we obtain

$$I_{\nu\nu}(\boldsymbol{\pi},\nu) = \sum_C\frac{n_C}{N}\mathcal{E}_\pi\left(\left(\sum_{x=2}^{m}\sum_{\{T_{j_1},\dots,T_{j_x}\}}\frac{w_{\{j_1,\dots,j_x\}|C,\alpha}}{\nu}-\frac{1}{D_C}\sum_{x=2}^{m}\sum_{\{T_{j_1},\dots,T_{j_x}\}\subseteq C}\sqrt[x]{\pi_{j_1}\dots\pi_{j_x}}\right)^2\right).$$

Using Equation 3.6, this simplifies to

$$I_{\nu\nu}(\boldsymbol{\pi},\nu) = \sum_C\frac{n_C}{N}\mathrm{Var}_\pi\left(\sum_{x=2}^{m}\sum_{\{T_{j_1},\dots,T_{j_x}\}\subseteq C}\frac{w_{\{j_1,\dots,j_x\}|C,\alpha}}{\nu}\right).$$

As in the other cases, we will expand the variance and substitute this expression and covariances of the single ws. After simplification, we obtain

$$\begin{aligned} I_{\nu\nu}(\boldsymbol{\pi},\nu) &= \sum_C\frac{n_C}{N\nu^2D_C^2}\times\Bigg(\sum_{x=2}^{m}\sum_{\{T_{j_1},\dots,T_{j_x}\}\subseteq C}\nu\sqrt[x]{\pi_{j_1}\dots\pi_{j_x}}(D_C-\nu\sqrt[x]{\pi_{j_1}\dots\pi_{j_x}}) \\ &\quad-\sum_{x=2}^{m}\sum_{u=2}^{m}\sum_{\substack{\{T_{j_1},\dots,T_{j_x}\}\\ \neq\{T_{j'_1},\dots,T_{j'_u}\}\subseteq C}}\nu^2\sqrt[x]{\pi_{j_1}\dots\pi_{j_x}}\sqrt[u]{\pi_{j'_1}\dots\pi_{j'_u}}\Bigg). \end{aligned}$$

As with the Davidson ties model in Chapter 2, our ultimate goal will be to estimate contrasts of the $\gamma_i = \ln(\pi_i)$, and not the π_is themselves. In order to achieve this, we will need to find the information matrix for the estimation of the entries in $\boldsymbol{\gamma}$ plus ν. This information matrix was introduced for the MNL model in Section 1.1, and was derived for the Davidson ties model in Chapter 2. We partition this information matrix in the same way as $I(\boldsymbol{\pi},\nu)$, giving

$$\Lambda(\boldsymbol{\pi},\nu) = \begin{bmatrix}\Lambda_{\gamma\gamma}(\boldsymbol{\pi},\nu) & \Lambda_{\nu\gamma}(\boldsymbol{\pi},\nu)\\ \Lambda_{\gamma\nu}(\boldsymbol{\pi},\nu) & \Lambda_{\nu\nu}(\boldsymbol{\pi},\nu)\end{bmatrix},$$

where $\Lambda_{\nu\gamma}(\boldsymbol{\pi},\nu) = \Lambda_{\gamma\nu}(\boldsymbol{\pi},\nu)^T$. As in the case of the Davidson ties models,

$$\Lambda(\boldsymbol{\pi},\nu) = PI(\boldsymbol{\pi},\nu)P^T$$

where

$$P = \begin{bmatrix} \pi_1 & 0 & \dots & 0 & 0 \\ 0 & \pi_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \pi_t & 0 \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix}.$$

If we look at each of the generic entries, we get

$$\begin{aligned} \Lambda_{\gamma\gamma}(\boldsymbol{\pi},\nu)_{ij} &= \pi_i\pi_j I_{\pi\pi}(\boldsymbol{\pi},\nu)_{ij} \\ \Lambda_{\gamma\gamma}(\boldsymbol{\pi},\nu)_{ii} &= \pi_i^2 I_{\pi\pi}(\boldsymbol{\pi},\nu)_{ii} \\ \Lambda_{\gamma\nu}(\boldsymbol{\pi},\nu)_{1i} &= \pi_i I_{\pi\nu}(\boldsymbol{\pi},\nu) \\ \Lambda_{\nu\nu}(\boldsymbol{\pi},\nu) &= I_{\nu\nu}(\boldsymbol{\pi},\nu). \end{aligned}$$

If we assume, as did Davidson [1970], the null hypothesis of equal merits for each of the items and that ν is left unspecified, then the expression for $\Lambda(\boldsymbol{\pi},\nu)$ simplifies greatly. That is, we assume that

$$\boldsymbol{\pi} = \boldsymbol{j} = \boldsymbol{\pi}_0,$$

and get

$$\begin{aligned} &\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0,\nu)_{ij} \\ &= \sum_{C|T_i,T_j\in C} \frac{n_C}{ND_C^2}\Bigg(-1 - \sum_{u=2}^{m}\sum_{\{T_i,T_{j'_2},\dots,T_{j'_u}\}\subseteq C}\frac{\nu}{u} - \sum_{x=2}^{m}\sum_{\{T_i,T_{j_2},\dots,T_{j_x}\}\subseteq C}\frac{\nu}{x} \\ &\qquad + \sum_{x=2}^{m}\sum_{\{T_i,T_j,T_{j_3},\dots,T_{j_x}\}}\frac{\nu(D_C-\nu)}{x^2} - \sum_{x=2}^{m}\sum_{u=2}^{m}\sum_{\substack{\{T_i,T_{j_2},\dots,T_{j_x}\}\\ \neq\{T_j,T_{j'_2},\dots,T_{j'_u}\}}}\frac{\nu^2}{xy}\Bigg) \\ &= \sum_{C|T_i,T_j\in C} \frac{n_C}{ND_C^2}\Bigg(-1 - \nu\bigg(\sum_{u=2}^{m}\frac{1}{u}\binom{m-1}{y-1} + \sum_{x=2}^{m}\frac{1}{x}\binom{m-1}{x-1}\bigg) + \nu(D_C-\nu)\sum_{x=2}^{m}\frac{1}{x^2}\binom{m-2}{x-2} \\ &\qquad -\nu^2\bigg(\sum_{x=2}^{m}\sum_{u=2}^{m}\frac{1}{x\times u}\binom{m-1}{x-1}\binom{m-1}{u-1} - \sum_{x=2}^{m}\frac{1}{x^2}\binom{m-2}{x-2}\bigg)\Bigg). \end{aligned}$$

If we use the identity

$$\frac{m}{x}\binom{m-1}{x-1} = \binom{m}{x}, \tag{3.7}$$

and simplify, then we obtain

$$\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0,\nu)_{ij} = \frac{\nu\sum_{x=2}^{m}\left(\left(\frac{(x-1)}{xm(m-1)} - \frac{1}{m^2}\right)\binom{m}{x}\right) - \frac{1}{m}}{N(m+\nu\sum_{x=2}^{m}\binom{m}{x})}\sum_{C|T_i,T_j\in C} n_C.$$

Next we look at the diagonal entries of $\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \nu)$. We have

$$\begin{aligned}
&\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \nu)_{ii} \\
&= \sum_{C|T_i\in C} \frac{n_C}{ND_C^2}\Bigg((D_C-1)+\sum_{x=2}^{m}\sum_{\{T_i,T_{j_2},\ldots,T_{j_x}\}\subseteq C}\frac{\nu(D_C-\nu)}{x^2}-2\sum_{x=2}^{m}\sum_{\{T_i,T_{j_2},\ldots,T_{j_x}\}\subseteq C}\frac{\nu}{x} \\
&\qquad -\sum_{x=2}^{m}\sum_{u=2}^{m}\sum_{\substack{\{T_i,T_{j_2},\ldots,T_{j_x}\}\\ \neq\{T_i,T_{j'_2},\ldots,T_{j'_u}\}\subseteq C}}\frac{\nu^2}{xu}\Bigg) \\
&= \sum_{C|T_i\in C} \frac{n_C}{ND_C^2}\Bigg((D_C-1)+\nu(D_C-\nu)\sum_{x=2}^{m}\frac{1}{x^2}\binom{m-1}{x-1}-2\nu\sum_{x=2}^{m}\frac{1}{x}\binom{m-1}{x-1} \\
&\qquad -\nu^2\Bigg(\sum_{x=2}^{m}\sum_{u=2}^{m}\frac{1}{xu}\binom{m-1}{x-1}\binom{m-1}{u-1}-\sum_{x=2}^{m}\frac{1}{x^2}\binom{m-1}{x-1}\Bigg)\Bigg),
\end{aligned}$$

which, by using Equation 3.7, can be simplified to

$$\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \nu)_{ii} = \frac{1-\frac{1}{m}+\nu\sum_{x=2}^{m}\left(\frac{1}{mx}-\frac{1}{m^2}\right)\binom{m}{x}}{N(m+\nu\sum_{x=2}^{m}\binom{m}{x})}\sum_{C|T_i\in C} n_C.$$

Next observe that

$$\begin{aligned}
&\Lambda_{\gamma\nu}(\boldsymbol{\pi}_0, \nu)_{i} \\
&= \sum_{i|T_i\in C} \frac{n_C}{N\nu D_C^2}\Bigg(-\sum_{u=2}^{m}\sum_{\{T_{j'_1},T_{j'_2},\ldots,T_{j'_u}\}\subseteq C}\nu+\sum_{x=2}^{m}\sum_{\{T_i,T_{j_2},\ldots,T_{j_x}\}\subseteq C}\frac{\nu(D_C-\nu)}{x} \\
&\qquad -\sum_{x=2}^{m}\sum_{u=2}^{m}\sum_{\substack{\{T_i,T_{j_2},\ldots,T_{j_x}\}\\ \neq\{T_{j'_1},T_{j'_2},\ldots,T_{j'_u}\}\subseteq C}}\frac{\nu^2}{x}\Bigg) \\
&= \sum_{C|T_i\in C} \frac{n_C}{ND_C^2}\Bigg(-\nu\sum_{x=2}^{m}\binom{m}{x}+\nu(D_C-\nu)\sum_{x=2}^{m}\frac{1}{x}\binom{m-1}{x-1} \\
&\qquad -\nu^2\Bigg(\sum_{x=2}^{m}\sum_{u=2}^{m}\frac{1}{xu}\binom{m-1}{x-1}\binom{m}{u}-\sum_{x=2}^{m}\frac{1}{x}\binom{m-1}{x-1}\Bigg)\Bigg),
\end{aligned}$$

and we use Equation 3.7 and simplify to give

$$\Lambda_{\gamma\nu}(\boldsymbol{\pi}_0, \nu)_{i} = \sum_{C|T_i\in C}\frac{n_C\left(-\sum_{x=2}^{m}\binom{m}{x}-\frac{\nu}{m}\left(\sum_{x=2}^{m}\binom{m}{x}\right)^2+\sum_{x=2}^{m}\binom{m}{x}+\frac{\nu}{m}\left(\sum_{x=2}^{m}\binom{m}{x}\right)^2\right)}{N\pi_i D_C^2},$$

which is clearly 0. Thus $\Lambda_{\gamma\nu}(\boldsymbol{\pi}_0, \nu) = \mathbf{0}$. Recall that these off–diagonal blocks were also zero for the Davidson ties model.

Finally, we have

$$\begin{aligned}
\Lambda_{\nu\nu}(\boldsymbol{\pi}_0, \nu) &= \sum_{C}\frac{n_C}{N\nu^2D_C^2}\times\Bigg(\sum_{x=2}^{m}\sum_{\{T_{j_1},\ldots,T_{j_x}\}\subseteq C}\nu(D_C-\nu)-\sum_{x=2}^{m}\sum_{u=2}^{m}\sum_{\substack{\{T_{j_1},\ldots,T_{j_x}\}\\ \neq\{T_{j'_1},\ldots,T_{j'_u}\}\subseteq C}}\nu^2\Bigg) \\
&= \sum_{C}\frac{n_C\left(\nu(D_C-\nu)\sum_{x=2}^{m}\binom{m}{x}-\nu^2\left(\sum_{x=2}^{m}\sum_{u=2}^{m}\binom{m}{x}\binom{m}{u}-\sum_{x=2}^{m}\binom{m}{x}\right)\right)}{N\nu^2D_C^2},
\end{aligned}$$

which, using Equation 3.7, simplifies to give

$$\Lambda_{\nu\nu}(\boldsymbol{\pi}_0, \nu) = \frac{m \sum_{x=2}^{m} \binom{m}{x}}{N\nu(m + \nu \sum_{x=2}^{m} \binom{m}{x})^2} \sum_C n_C.$$

Noting that $N = \sum_C n_C$, this will simplify further to give

$$\Lambda_{\nu\nu}(\boldsymbol{\pi}_0, \nu) = \frac{m \sum_{x=2}^{m} \binom{m}{x}}{\nu(m + \nu \sum_{x=2}^{m} \binom{m}{x})^2}.$$

This is a function of m and ν only and hence we refer to this entry as $\Lambda_{\nu\nu}(m, \nu)$.

Now let us return to our example and derive the $\Lambda(\boldsymbol{\pi}_0, \nu)$ matrix.

■ **EXAMPLE 3.2.1.**

Recall the experiment introduced in Example 3.1.1. The information matrix for the estimation of $\boldsymbol{\gamma}$ and ν under the null hypothesis of equal merits is

$$\Lambda(\boldsymbol{\pi}_0, \nu) = \frac{(4+\nu)}{24(3+4\nu)} \begin{bmatrix} 3 & -1 & -1 & -1 & 0 \\ -1 & 3 & -1 & -1 & 0 \\ -1 & -1 & 3 & -1 & 0 \\ -1 & -1 & -1 & 3 & 0 \\ 0 & 0 & 0 & 0 & \frac{288}{\nu(3+4\nu)(4+\nu)} \end{bmatrix}.$$

□

Now that we have an information matrix for the entries in $\boldsymbol{\gamma}$ and for ν we are able to construct an information matrix for the estimation of ν and contrasts of the entries in $\boldsymbol{\gamma}$. This matrix will then be used to develop some results about the optimal design for sets of contrasts when this model is appropriate.

3.3 Representing options using k attributes

In this section we consider the construction of the information matrix for the estimation of ν and contrasts of the entries in $\boldsymbol{\gamma}$. In particular, we are interested in the estimation of the contrasts of the entries in $\boldsymbol{\gamma}$ that relate to the main effects and interaction effects of the attributes, as introduced in Section 1.B.

We begin by constructing a matrix B that contains coefficients of linear combinations of the entries in $\boldsymbol{\gamma}$ and ν. We assume that any interaction between the entries in $\boldsymbol{\gamma}$ and ν are not of interest. This allows us to partition B into two non–zero blocks. The first block is B_γ, which contains contrasts of the entries in $\boldsymbol{\gamma}$. The other block will equal 1. Appendix 3.B shows that this will not violate any of the assumptions forced on the information matrix by El-Helbawy and Bradley [1978]. The partitioned B matrix can be expressed as

$$B = \begin{bmatrix} B_\gamma & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix}.$$

Then the information matrix for the estimation of ν and the contrasts in $B_\gamma\boldsymbol{\gamma}$, under the null hypothesis of equal merits, is

$$\begin{aligned} C(\boldsymbol{\pi}_0, \nu) &= B\Lambda(\boldsymbol{\pi}_0, \nu)B^T \\ &= \begin{bmatrix} B_\gamma & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \nu) & \mathbf{0} \\ \mathbf{0} & \Lambda_{\nu\nu}(m, \nu) \end{bmatrix} \begin{bmatrix} B_\gamma^T & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \\ &= \begin{bmatrix} B_\gamma \Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \nu) B_\gamma^T & \mathbf{0} \\ \mathbf{0} & \Lambda_{\nu\nu}(m, \nu) \end{bmatrix}. \end{aligned}$$

Now let us reconsider the experiment introduced in Example 3.1.1, calculating the information matrix for the estimation of ν and the main effects.

■ **EXAMPLE 3.3.1.**

Consider the experiment introduced in Example 3.1.1 and the design introduced in Example 3.1.3 to estimate main effects and ν. The B for the estimation of ν and the main effects is

$$\begin{aligned} B &= \begin{bmatrix} B_\gamma & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix} -1 & -1 & 1 & 1 & 0 \\ -1 & 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}, \end{aligned}$$

where B_γ is a 2×4 matrix of contrast coefficients. Then the information matrix for the estimation of ν and the main effects is

$$C(\boldsymbol{\pi}_0,\nu) = \begin{bmatrix} \frac{4+\nu}{6(3+4\nu)} & 0 & 0 \\ 0 & \frac{4+\nu}{6(3+4\nu)} & 0 \\ 0 & 0 & \frac{12}{\nu(3+4\nu)^2} \end{bmatrix},$$

Since the information matrix is diagonal, we can estimate ν and the main effects independently when using this design. □

Now that we have the information matrix for the effects that are of interest, we can determine the optimal designs for the estimation of these effects.

3.4 Optimal designs for the estimation of effects in the generalised Davidson ties model

In this section we compare the information matrices for the estimation of a given set of contrasts when the generalised Davidson ties model is used to that obtained when the MNL model is used. We then show that designs that are optimal for the MNL model are also optimal for the estimation of ν and the same set of contrasts for the generalised Davidson ties model.

Recall from Section 1.1 that the generic entries for $\Lambda(\boldsymbol{\pi}_0)$ when the MNL model is used, denoted here as $\Lambda(\boldsymbol{\pi}_0)_{\text{MNL}}$, are

$$(\Lambda(\boldsymbol{\pi}_0)_{\text{MNL}})_{ij} = \frac{-1}{m^2 N}\sum_{C|T_i,T_j\in C} n_C,$$

and

$$(\Lambda(\boldsymbol{\pi}_0)_{\text{MNL}})_{ii} = \frac{m-1}{m^2 N}\sum_{C|T_i\in C} n_C.$$

If we compare this to the first block of $\Lambda(\boldsymbol{\pi}_0,\nu)$ when the generalised Davidson ties model is used, denoted here as $(\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0,\nu)_{\text{DAV}})$, then we obtain

$$\frac{1}{N}\sum_{C|T_i,T_j\in C} n_C = -m^2(\Lambda(\boldsymbol{\pi}_0)_{\text{MNL}})_{ij} = \frac{(m-1)(m+\nu\sum_{x=2}^m\binom{m}{x})}{\nu\sum_{x=2}^m((\frac{1}{m^2}-\frac{1}{mx})\binom{m}{x})}(\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0,\nu)_{\text{DAV}})_{ij},$$

which simplifies to give

$$(\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0,\nu)_{\text{DAV}})_{ij} = Q(m,\nu) \times (\Lambda(\boldsymbol{\pi}_0)_{\text{MNL}})_{ij},$$

where

$$Q(m,\nu) = \frac{m + \nu \sum_{x=2}^{m} (\frac{m-x}{x(m-1)} \binom{m}{x})}{m + \nu \sum_{x=2}^{m} \binom{m}{x}}.$$

Similarly, we find that

$$\frac{1}{N} \sum_{C|T_i \in C} n_C = \frac{m^2}{m-1} (\Lambda(\boldsymbol{\pi}_0)_{\text{MNL}})_{ii} = \frac{(m + \nu \sum_{x=2}^{m} \binom{m}{x})}{\frac{m-1}{m} + \nu \sum_{x=2}^{m} ((\frac{1}{mx} - \frac{1}{m^2}) \binom{m}{x})} (\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0,\nu)_{\text{DAV}})_{ii},$$

which simplifies to give

$$(\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0,\nu)_{\text{DAV}})_{ij} = Q(m,\nu) \times (\Lambda(\boldsymbol{\pi}_0)_{\text{MNL}})_{ij}.$$

Thus we see that

$$\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0,\nu)_{\text{DAV}} = Q(m,\nu) \times \Lambda(\boldsymbol{\pi}_0)_{\text{MNL}},$$

which gives

$$\Lambda(\boldsymbol{\pi}_0,\nu)_{\text{DAV}} = \begin{pmatrix} Q(m,\nu) \times \Lambda(\boldsymbol{\pi}_0)_{\text{MNL}} & 0 \\ 0 & \Lambda_{\nu\nu}(m,\nu) \end{pmatrix}.$$

Since we can express the information matrix for the estimation of ν and contrasts of the entries in $\boldsymbol{\gamma}$ when the generalised Davidson ties model is used in terms of the information matrix for the estimation of the entries in $\boldsymbol{\gamma}$ when the MNL model is used, we may now look at comparing optimality criteria for designs using these two models. We will use the D–optimality criterion as defined in Section 1.3.1.

■ **THEOREM 3.4.1.**
For a set of contrasts of the entries in $\boldsymbol{\gamma}$ and a constant but unknown ν, the D–optimal design for the estimation of the contrasts of $\boldsymbol{\gamma}$ over a set of competing designs $\mathfrak{X}$ when the MNL model is used will also be D–optimal for the estimation of the same contrasts and ν over the same set of competing designs for the estimation of the generalised Davidson ties model. □

Proof. We begin by letting B be a block diagonal matrix

$$B = \begin{bmatrix} B_\gamma & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix},$$

where B_γ is a $p \times t$ matrix containing the coefficients of the contrasts of $\boldsymbol{\gamma} = \ln(\boldsymbol{\pi})$ that are of interest. The contrasts in $B_\gamma \boldsymbol{\gamma}$ are to be estimated in both models. Then the information matrix for the estimation of the contrasts in $B_\gamma \boldsymbol{\gamma}$ as well as ν when the generalised Davidson model is used is

$$C(\boldsymbol{\pi}_0,\nu)_{\text{DAV}} = \begin{bmatrix} B_\gamma \Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0,\nu) B_\gamma^T & \mathbf{0} \\ \mathbf{0} & \Lambda_{\nu\nu}(m,\nu) \end{bmatrix}.$$

However, we have shown that

$$\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0,\nu)_{\text{DAV}} = Q(m,\nu) \times \Lambda(\boldsymbol{\pi}_0)_{\text{MNL}},$$

so by substituting this into the expression for $C(\boldsymbol{\pi}_0,\nu)_{\text{DAV}}$ we obtain

$$C(\boldsymbol{\pi}_0,\nu)_{\text{DAV}} = \begin{bmatrix} Q(m,\nu) B_\gamma \Lambda(\boldsymbol{\pi}_0)_{\text{MNL}} B_\gamma^T & \mathbf{0} \\ \mathbf{0} & \Lambda_{\nu\nu}(m,\nu) \end{bmatrix}.$$

We notice that the information matrix for the estimation of the set of effects in $B_\gamma \boldsymbol{\gamma}$ when the MNL model is used is

$$C(\boldsymbol{\pi}_0)_{\text{MNL}} = B_\gamma \Lambda(\boldsymbol{\pi}_0)_{\text{MNL}} B_\gamma^T.$$

That is,

$$C(\boldsymbol{\pi}_0, \nu)_{\text{DAV}} = \begin{pmatrix} Q(m,\nu)C(\boldsymbol{\pi}_0)_{\text{MNL}} & 0 \\ 0 & \Lambda_{\nu\nu}(m,\nu) \end{pmatrix}.$$

Then

$$\det(C(\boldsymbol{\pi}_0, \nu)_{\text{DAV}}) = (Q(m,\nu))^p \times \Lambda_{\nu\nu}(m,\nu) \times \det(C(\boldsymbol{\pi}_0)_{\text{MNL}}).$$

Since

$$\det((C(\boldsymbol{\pi}_0)_{\text{MNL}})_{\xi_{OPT}}) \geq \det((C(\boldsymbol{\pi}_0)_{\text{MNL}})_\xi)$$

for all $\xi \in \mathfrak{X}$, the relative efficiency of a generic design compared to the design ξ_{OPT}, which is optimal for the estimation of the set of contrasts using the MNL model, when the generalised Davidson ties model is used, is

$$\begin{aligned} D_{\text{eff}}(\xi, \xi_{\text{OPT}}) &= \left(\frac{\det(C_{\xi,\text{DAV}})}{\det(C_{\xi_{\text{OPT}},\text{DAV}})}\right)^{\frac{1}{p+1}} \\ &= \left(\frac{(Q(m,\nu))^p \times \Lambda_\nu(m,\nu) \times \det(C_{\xi,\text{MNL}})}{(Q(m,\nu))^p \times \Lambda_\nu(m,\nu) \times \det(C_{\xi_{\text{OPT}},\text{MNL}})}\right)^{\frac{1}{p+1}} \\ &= \left(\frac{\det(C_{\xi,\text{MNL}})}{\det(C_{\xi_{\text{OPT}},\text{MNL}})}\right)^{\frac{1}{p+1}} \\ &\leq 1, \end{aligned}$$

for all $\xi \in \mathfrak{X}$. Therefore, by the definition of D–optimality, the design ξ_{OPT} is also optimal for the estimation of ν and the set of contrasts in $B_\gamma \boldsymbol{\gamma}$ when the generalised Davidson ties model is used. □

We now consider an example of the relationship between the two models, and also compare some designs.

■ **EXAMPLE 3.4.1.**

Recall the experiment and design introduced in Example 3.1.3. In Example 3.3.1 we found the information matrix for the estimation of ν and the main effects when the generalised Davidson ties model is used. Now we will find the information matrix for the estimation of main effects only for the same design where the MNL model is used. The contrast matrix for the estimation of main effects is

$$B = \frac{1}{2} \begin{bmatrix} -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \end{bmatrix}.$$

From Burgess and Street [2003] we know that the information matrix for the estimation of the entries in $\boldsymbol{\gamma}$ using the MNL model under the assumption of the null hypothesis of equal merits is

$$\Lambda(\boldsymbol{\pi}_0, \nu)_{\text{MNL}} = \frac{1}{8} \begin{bmatrix} 3 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ -1 & -1 & -1 & 3 \end{bmatrix}.$$

Notice that

$$\Lambda_\gamma(\boldsymbol{\pi}_0,\nu)_{\text{DAV}} = \frac{3(4+\nu)}{4(3+4\nu)}\Lambda(\boldsymbol{\pi}_0)_{\text{MNL}}.$$

It follows that the information matrix for the estimation of main effects only using the MNL model is

$$C(\boldsymbol{\pi}_0)_{\text{MNL}} = \begin{bmatrix} \frac{2}{9} & 0 \\ 0 & \frac{2}{9} \end{bmatrix},$$

so we see that

$$C(\boldsymbol{\pi}_0,\nu)_{\text{DAV}} = \begin{bmatrix} \frac{3(4+\nu)}{4(3+4\nu)}C(\boldsymbol{\pi}_0)_{\text{MNL}} & 0 \\ 0 & \frac{12}{\nu(3+4\nu)^2} \end{bmatrix}.$$

Taking determinants of both $C(\boldsymbol{\pi}_0,\nu)_{\text{DAV}}$ and $C(\boldsymbol{\pi}_0)_{\text{MNL}}$ gives

$$\det(C(\boldsymbol{\pi}_0,\nu)_{\text{DAV}}) = \frac{(4+\nu)^2}{3\nu(3+4\nu)^6},$$

and

$$\det(C(\boldsymbol{\pi}_0)_{\text{MNL}}) = \frac{4}{81}.$$

We are estimating $p=2$ contrasts on $\boldsymbol{\gamma}$, so

$$\begin{aligned}(Q(m,\nu))^p \times \Lambda_\nu(m,\nu) \times \det(C(\boldsymbol{\pi}_0)_{\text{MNL}}) &= \frac{12\times(3+\frac{3}{4}\nu)^2}{\nu(3+4\nu)^6} \times \frac{4}{81} \\ &= \frac{(4+\nu)^2}{3\nu(3+4\nu)^6} \\ &= \det(C(\boldsymbol{\pi}_0,\nu)_{\text{DAV}}),\end{aligned}$$

which is consistent with the findings in Theorem 3.4.1. □

We can use this theorem to apply some of the known results for the MNL model to the generalised Davidson ties model. First, we consider an extension of the theorem for the estimation of main effects in a 2^k model presented in Burgess and Street [2003] .

■ **COROLLARY 3.4.2.**

The D–optimal design for the estimation of ν and the main effects, when all other effects are assumed to be zero, using the generalised Davidson ties model is given by the choice sets in which, for each $\boldsymbol{v}_i$ present,

$$\sum_{i=1}^{m-1}\sum_{j=i+1}^{m} d_{ij} = \begin{cases} \frac{(m^2-1)k}{4}, & m \text{ odd}, \\ \frac{m^2k}{4}, & m \text{ even}, \end{cases}$$

and there is at least one $\boldsymbol{v}_j$ with a non–zero a_j; that is, the choice set is non–empty. □

Proof. By Theorem 1.3.4, the design described in the statement of the theorem is optimal for the estimation of main effects only when the MNL model is used. Then it follows from Theorem 3.4.1 that this design is also optimal for the estimation of ν and the main effects when the generalised Davidson ties model is used for a 2^k factorial experiment. □

Now let us use this theorem to find an optimal design for the estimation of ν and the main effects for our example.

■ **EXAMPLE 3.4.2.**

Consider the 2^2 experiment introduced in Example 3.1.1. In this experiment $\frac{m^2-1}{4} = 4$, so the D–optimal design is given by the choice sets with difference vectors whose entries sum to 4. The only distinct difference vector that doesn't force repeated items in a choice set is $\boldsymbol{v} = (01, 10, 11)$, giving the design in Table 3.2. □

Option 1	Option 2	Option 3
0 0	0 1	1 0
0 1	0 0	1 1
1 0	1 1	0 0
1 1	1 0	0 1

Table 3.2: The 2^2 design that is optimal for estimating ν and the main effects.

Now we consider an extension to the theorem for the estimation of main effects plus two–factor interactions in a 2^k factorial presented in Burgess and Street [2003].

■ **COROLLARY 3.4.3.**

The D–optimal design for the estimation of ν and the main effects plus two–factor interactions, using the generalised Davidson ties model, when all other effects are assumed to be zero, is given by

$$y_i = \begin{cases} \frac{m(m-1)}{2^k}\binom{k+1}{k/2}^{-1}, & k \text{ even and } i = k/2, k/2+1, \\ \frac{m(m-1)}{2^k}\binom{k}{(k+1)/2}^{-1}, & k \text{ odd and } i = (k+1)/2, \\ 0, & \text{otherwise}, \end{cases}$$

when this results in non–zero y_is that correspond to difference vectors that actually exist. □

Proof. By Theorem 1.3.5, the design described in the statement of the theorem is optimal for the estimation of main effects plus two–factor interactions when the MNL model is used. Then it follows from Theorem 3.4.1 that this design is also optimal for the estimation of ν and the main effects plus two–factor interactions when the generalised Davidson ties model is used for a 2^k factorial experiment. □

Now let us use this theorem to find an optimal design for the estimation of main effects plus two–factor interactions and ν for the experiment in our example.

■ **EXAMPLE 3.4.3.**

Consider the 2^2 experiment introduced in Example 3.1.1. To obtain the D–optimal design for the estimation of ν and the main effects, two–factor interactions using the generalised Davidson ties model we need

$$y_i = \begin{cases} \frac{1}{2}, & \text{if } i = 1 \text{ or } 2, \\ 0, & \text{otherwise}. \end{cases}$$

This gives the design in Table 3.3. □

Finally, we look at an extension for the theorem for the estimation of main effects in an experiment with asymmetric attributes presented in Burgess and Street [2005] to include the estimation of ν as well.

■ **COROLLARY 3.4.4.**

Let F be the complete factorial for k attributes where the q^{th} attribute has ℓ_q levels. Suppose that we choose a set of m generators $G = \{\boldsymbol{g}_1 = \mathbf{0}, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_m\}$ such that $\boldsymbol{g}_i \neq \boldsymbol{g}_j$ for $i \neq j$. Suppose that $g_i = (g_{i1}, g_{i2}, \ldots, g_{ik})$ for $i = 1, \ldots, m$, and suppose that the multiset of differences for attribute q, $\{\pm(g_{i_1q} - g_{i_2q}) | 1 \leq i_1, i_2 \leq m, i_1 \neq i_2\}$, contains each non–zero difference modulo

Option 1	Option 2	Option 3
0 0	0 1	1 0
0 1	0 0	1 1
1 0	1 1	0 0
1 1	1 0	0 1

Table 3.3: The optimal 2^2 design for estimating ν and the main effects plus two–factor interactions.

ℓ_q equally often. Then the choice sets given by the rows of $F + \boldsymbol{g}_1, F + \boldsymbol{g}_2, \ldots, F + \boldsymbol{g}_m$ for one or more sets of generators, are optimal for the estimation of ν and the main effects using the generalised Davidson ties model, provided that there are as few zero differences as possible. □

Proof. By Theorem 1.3.6, the design described in the statement of the theorem is optimal for the estimation of main effects when the MNL model is used. Then it follows from Theorem 3.4.1 that this design is also optimal for the estimation of ν and the main effects when the generalised Davidson ties model is used. □

Let us consider an example of how this theorem can be used to find optimal designs.

■ **EXAMPLE 3.4.4.**
Consider the 3^2 experiment introduced in Example 2.4.4. Now we present triples to the respondent. If we choose the set of generators $\boldsymbol{g} = (00, 11, 22)$ then there are no zero differences and each difference modulo 3 appears equally often. Then the design in Table 3.4, produced from this set of generators and removing any repeated choice sets, is D–optimal for the estimation of ν and the main effects when the generalised Davidson ties model is used. □

Option 1	Option 2	Option 3
0 0	1 1	2 2
1 2	2 0	0 1
2 1	0 2	1 0

Table 3.4: The 3^2 design that is optimal for estimating main effects and ν.

3.5 Simulations of the generalised Davidson ties model

In this section we consider the performance of the generalised Davidson ties model under various model assumptions by carrying out a number of simulation studies. We assume that $k = 2$, $\ell_1 = \ell_2 = 2$ and $m = 3$ throughout. We consider two sets of values for the parameters. In the first we assume that both main effects parameters, τ_1 and τ_2, are equal to 0 and the ties parameter $\nu = 0.5$, and in the second set we assume that $\tau_1 = 1$ and $\tau_2 = -1$ but $\nu = 0.5$ still.

We find the locally optimal design for each set of values and compare the performance of each design with both sets of parameter values. The design in Table 3.5 is optimal for the estimation of the main effects of the attributes plus the ties parameter when $\tau_1 = \tau_2 = 0$, and $\nu = 0.5$, by Corollary 3.4.2. By an exhaustive search of the $2^4 - 1 = 15$ possible designs, the design in Table 3.6 is optimal for the estimation of the main effects of the attributes plus the ties parameter when $\tau_1 = 1$, $\tau_2 = -1$, and $\nu = 0.5$. This exhaustive search is illustrated in Figure 3.1, where the x–coordinate corresponds to the design index, and the y–coordinate is the determinant of the information matrix for that design when $\tau_1 = 1, \tau_2 = -1$, and $\nu = 0.5$. The determinants of the information matrix for the designs in Tables 3.5 and 3.6 are labelled in Figure 3.1.

We first assume that $\tau_1 = \tau_2 = 0$, and $\nu = 0.5$ and compare the simulated distributions of the parameter estimates when the designs in Tables 3.5 and 3.6 are used in turn. Each simulation is modelled using the simulated responses from 150 respondents, and each boxplot displays the distribution of the estimates from 1000 such simulations. Figures 3.2(a) and (b) show the distributions of the parameter estimates when the designs in Tables 3.5 and 3.6, respectively, are used. Summary statistics for both simulations are provided in Table 3.7. We see from each of the figures that the distributions of the parameter estimates are symmetrically distributed. As expected, the variance of the parameter estimates for the design in Table 3.5 is smaller than that of the design in Table 3.6, illustrating the efficiency of the former design.

We now consider the performance of these two designs when $\tau_1 = 1, \tau_2 = -1$, and $\nu = 0.5$. Figures 3.3(a) and (b) show the distributions of the parameter estimates when the designs in Tables 3.5 and 3.6, respectively, are used. Summary statistics for both simulations are provided in Table 3.8. We see that, for both designs, the distribution of the parameter estimates seem to be unbiased and close to symmetric. The difference between the variances arising from the two designs is smaller in this case than when $\tau_1 = \tau_2 = 0$. The selection probabilities when $\tau_1 = 1, \tau_2 = -1$, and $\nu = 0.5$ for the design in Table 3.5 are given in Table 3.9.

Next, we simulate the effect of changing the magnitude of the ties parameter on the distribution of the parameter estimates when we let $\tau_1 = 1$ and $\tau_2 = -1$, and use the design in Table 3.6. Figures 3.4(a) and (b) give the simulated distributions of the parameter estimates when $\nu = 0.25$, and $\nu = 1$, respectively. Summary statistics for both simulations are provided in Table 3.10. Again, we see that the τ estimates are unbiased and symmetrically distributed. The variance of the estimate of ν increases as the magnitude of ν increases, while there is little difference between the variances of τ_1 and τ_2.

We now compare the ability of four different designs to estimate the main effects plus the two–factor interaction of the attributes and ν. The first two designs are those in Tables 3.5

Option 1	**Option 2**	**Option 3**
0 0	0 1	1 0
0 1	0 0	1 1
1 0	1 1	0 0
1 1	1 0	0 1

Table 3.5: Optimal design for estimating main effects and ν when $\tau_1 = \tau_2 = 0$, and $\nu = 0.5$.

Option 1	Option 2	Option 3
0 0	1 0	1 1
0 1	1 0	1 1

Table 3.6: Optimal design for estimating main effects and ν when $\tau_1 = 1, \tau_2 = -1$, and $\nu = 0.5$.

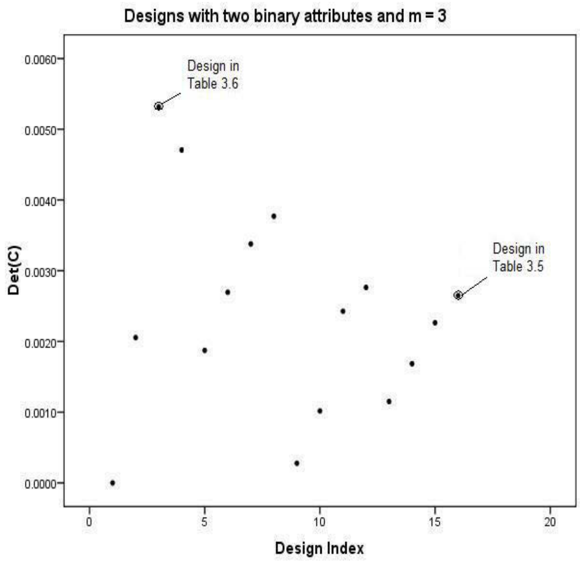

Figure 3.1: Exhaustive search for optimal design $\tau_1 = 1, \tau_2 = -1$, and $\nu = 0.5$.

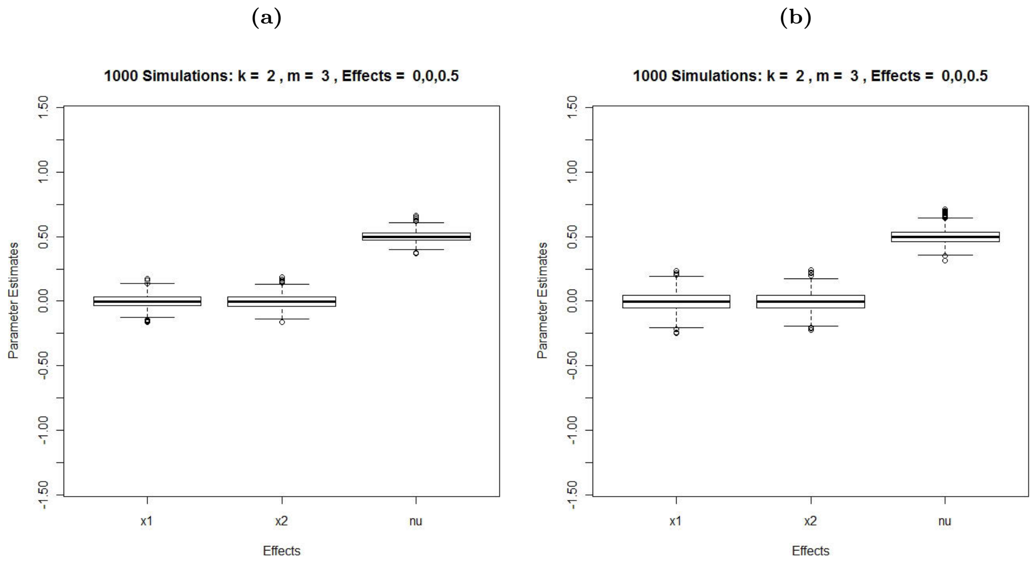

Figure 3.2: Simulation of Davidson ties model $\tau_1 = \tau_2 = 0$, and $\nu = 0.5$.

Parameter	Simulated Mean (Standard Error)	Theoretical Variance	Simulated Variance	Simulated Skewness (Standard Error)
Design in Table 3.5				
τ_1	−0.00088(0.00165)	0.00278	0.00271	0.02920(0.07734)
τ_2	−0.00244(0.00165)	0.00278	0.00273	0.13122(0.07734)
ν	0.50130(0.00134)	0.00694	0.00180	0.39140(0.07734)
Design in Table 3.6				
τ_1	−0.00445(0.00237)	0.00556	0.00564	−0.04829(0.07734)
τ_2	−0.00303(0.00232)	0.00556	0.00536	−0.10379(0.07734)
ν	0.50164(0.00184)	0.01389	0.00339	0.29694(0.07734)

Table 3.7: Summary statistics for $\tau_1 = \tau_2 = 0$, and $\nu = 0.5$.

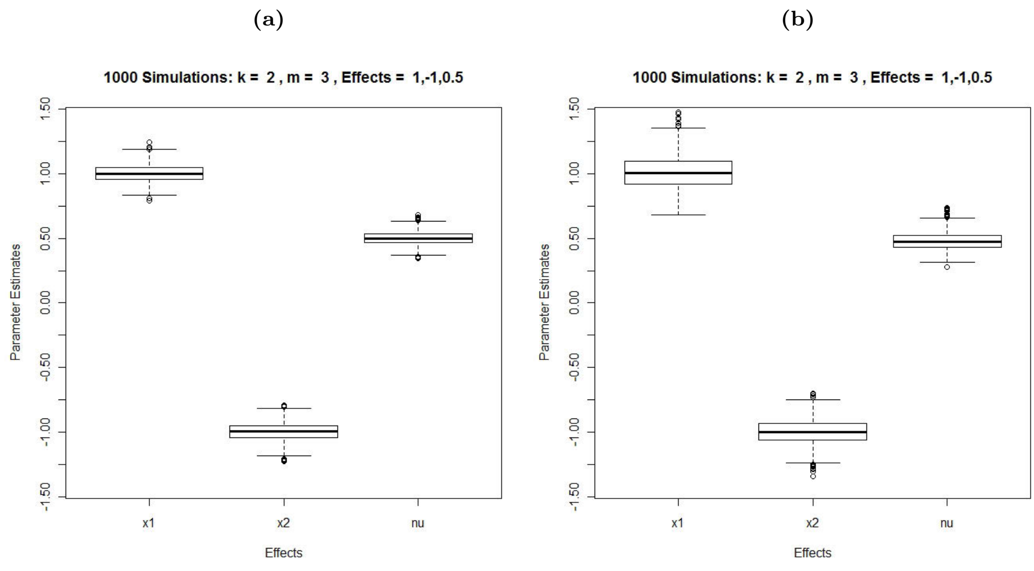

Figure 3.3: Simulation of Davidson ties model $\tau_1 = 1$, $\tau_2 = -1$, and $\nu = 0.5$.

Parameter	Simulated Mean (Standard Error)	Theoretical Variance	Simulated Variance	Simulated Skewness (Standard Error)
Design in Table 3.5				
τ_1	1.00248(0.00216)	0.00320	0.00465	0.20895(0.07734)
τ_2	−0.99894(0.00226)	0.00320	0.00510	−0.21943(0.07734)
ν	0.50071(0.00156)	0.00702	0.00244	0.19319(0.07734)
Design in Table 3.6				
τ_1	1.01053(0.00417)	0.00545	0.01736	0.36056(0.07734)
τ_2	−0.99980(0.00307)	0.00544	0.00941	−0.18175(0.07734)
ν	0.48061(0.00229)	0.01406	0.00527	0.51062(0.07734)

Table 3.8: Summary statistics for $\tau_1 = 1$, $\tau_2 = -1$, and $\nu = 0.5$.

Choice Set	$\mathbf{T_1}$	$\mathbf{T_2}$	$\mathbf{T_3}$	$\{\mathbf{T_1}, \mathbf{T_2}\}$	$\{\mathbf{T_1}, \mathbf{T_3}\}$	$\{\mathbf{T_2}, \mathbf{T_3}\}$	$\{\mathbf{T_1}, \mathbf{T_2}, \mathbf{T_3}\}$
$\{00, 01, 10\}$	0.090	0.012	0.668	0.017	0.123	0.045	0.045
$\{01, 00, 11\}$	0.042	0.307	0.307	0.056	0.056	0.153	0.079
$\{10, 11, 00\}$	0.544	0.074	0.074	0.100	0.100	0.037	0.072
$\{11, 10, 01\}$	0.090	0.668	0.012	0.123	0.017	0.045	0.045

Table 3.9: Selection probabilities when $\tau_1 = 1$, $\tau_2 = -1$, and $\nu = 0.5$.

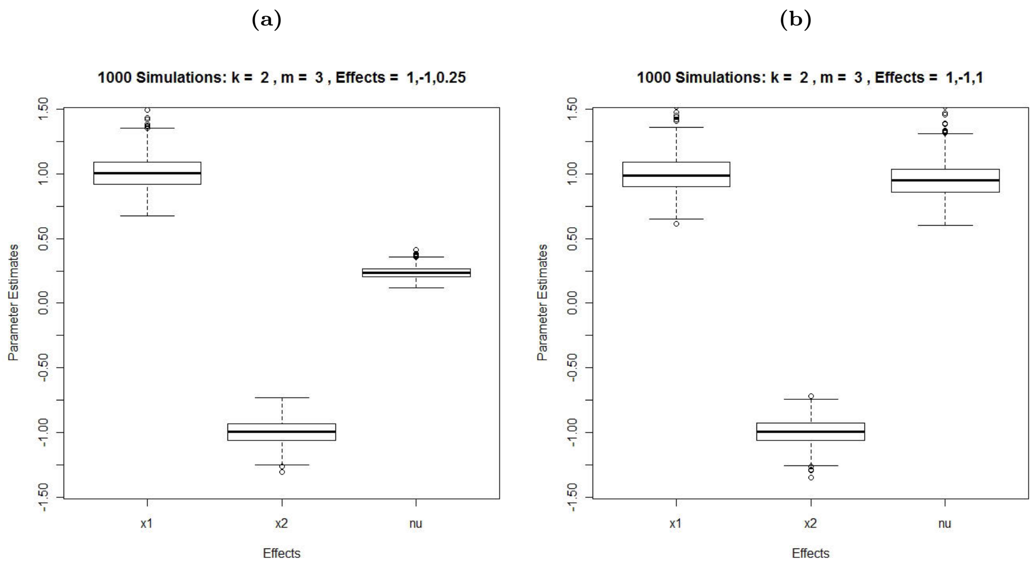

Figure 3.4: Simulation of Davidson ties model $\tau_1 = 1$, $\tau_2 = 0.5$, and (a) $\nu = 0.25$ and (b) $\nu = 1$.

	Parameter	Simulated Mean (Standard Error)	Theoretical Variance	Simulated Variance	Simulated Skewness (Standard Error)
(a)	τ_1	1.01026(0.00415)	0.00468	0.01724	0.36999(0.07734)
	τ_2	−1.00014(0.00299)	0.00468	0.00895	−0.18975(0.07734)
	ν	0.23830(0.00138)	0.01822	0.00191	0.41760(0.07734)
(b)	τ_1	0.99918(0.00443)	0.00680	0.01962	0.28574(0.07734)
	τ_2	−0.99508(0.00319)	0.00678	0.01018	−0.20531(0.07734)
	ν	0.95654(0.00436)	0.01358	0.01905	0.49900(0.07734)

Table 3.10: Summary statistics for $\tau_1 = 1$, $\tau_2 = -1$, and (a) $\nu = 0.25$ and (b) $\nu = 1$.

and 3.6. The third design is shown in Table 3.11, and is optimal for the estimation of the main effects plus the two–factor interaction of the attributes and ν when $\tau_1 = \tau_2 = \tau_{12} = 0$, and $\nu = 0.5$, by Corollary 3.4.3. The final design, shown in Table 3.12, is locally optimal for the estimation of the main effects plus two–factor interaction of the attributes and ν when $\tau_1 = 1, \tau_2 = -1, \tau_{12} = -0.25$, and $\nu = 0.5$ by an exhaustive search.

We first consider the case where there is no significant interaction effect. We let $\tau_1 = 1$, $\tau_2 = -1$, $\tau_{12} = 0$, and $\nu = 0.5$. Then Figures 3.5(a), (b), (c), and (d) give the simulated distributions of the parameter estimates when the designs in Table 3.5, 3.6, 3.11, and 3.12 are used. Summary statistics for all four of the simulations are provided in Table 3.13.

The design in Table 3.11 gives parameter estimates with the smallest variance, and are also unbiased and symmetrically distributed. The designs in Tables 3.5 and 3.12 also give unbiased and symmetric parameter estimates, but with a larger variance than those form the design in Table 3.11. This is expected, since there are three times as many choice sets in the design in Table 3.11. The design in Table 3.6 gives parameter estimates that are slightly biased towards 0, skewed, and with the largest variance of the four designs.

Now we consider the case where there is a significant interaction effect. Suppose that $\tau_1 = 1$, $\tau_2 = -1$, $\tau_{12} = -0.25$, and $\nu = 0.5$. Then Figures 3.6(a), (b), (c), and (d) give the simulated distributions of the parameter estimates when the designs in Table 3.5, Table 3.6, Table 3.11, and Table 3.12 are used. Summary statistics for all four of the simulations are provided in Table 3.14.

Again we see that the design in Table 3.11 gives parameter estimates with the smallest variance, and are also unbiased and symmetrically distributed. The designs in Tables 3.5 and 3.12 once again give unbiased and symmetric parameter estimates, but with a larger variance than those form the design in Table 3.11. The design in Table 3.6 once again gives parameter estimates that are slightly biased towards 0, skewed, and with the largest variance of the four designs.

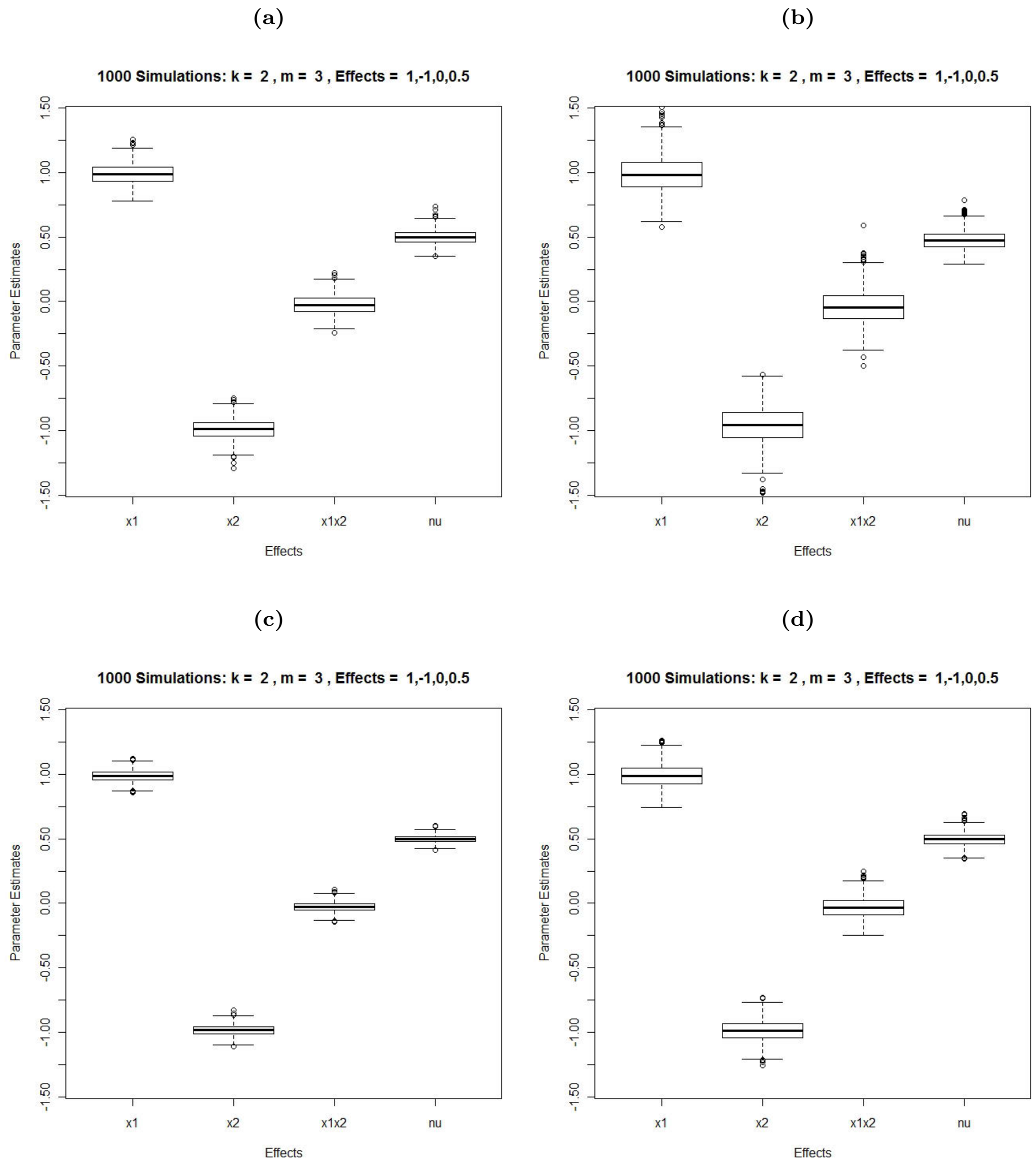

Figure 3.5: Simulation: estimating main effects and ν, designs in (a) Table 3.5, (b) Table 3.6, (c) Table 3.11, and (d) Table 3.12.

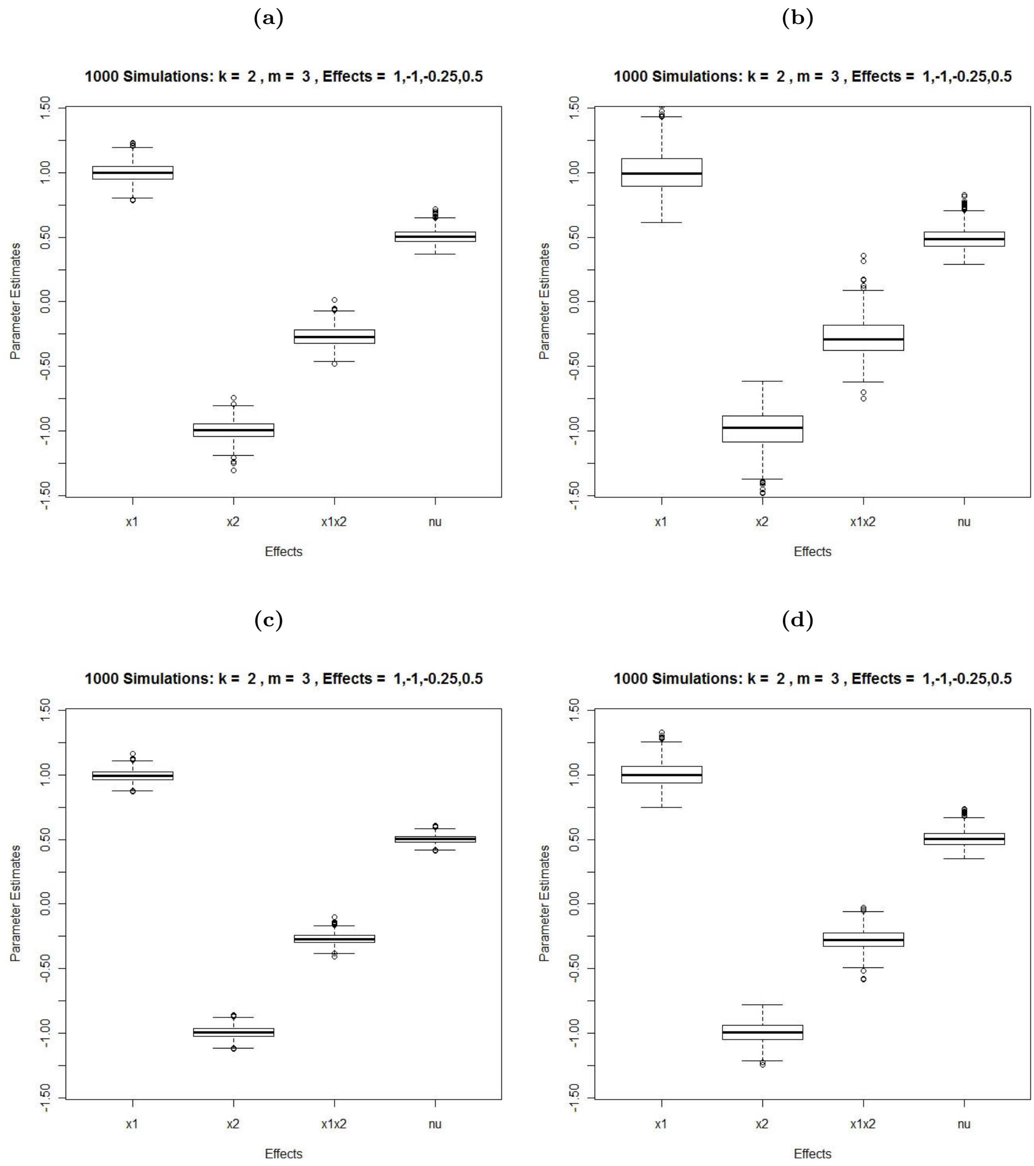

Figure 3.6: Simulation: estimating main effects, two–factor interactions, and ν, designs in (a) Table 3.5, (b) Table 3.6, (c) Table 3.11, and (d) Table 3.12.

Option 1	**Option 2**	**Option 3**
0 0	0 1	1 0
0 1	0 0	1 1
1 0	1 1	0 0
1 1	1 0	0 1
0 0	1 0	1 1
0 1	1 1	1 0
1 0	0 0	0 1
1 1	0 1	0 0
0 0	0 1	1 1
0 1	0 0	1 0
1 0	1 1	0 1
1 1	1 0	0 0

Table 3.11: Optimal design for estimating main effects, two–factor interactions and ν when $\tau_1 = \tau_2 = \tau_{12} = 0$, and $\nu = 0.5$.

Option 1	**Option 2**	**Option 3**
0 0	0 1	1 0
0 0	0 1	1 1
0 0	1 1	1 0
1 1	1 0	0 1

Table 3.12: Optimal design for estimating main effects, two–factor interactions and ν when $\tau_1 = 1, \tau_2 = -1, \tau_{12} = -0.25$, and $\nu = 0.5$.

Parameter	Simulated Mean (Standard Error)	Theoretical Variance	Simulated Variance	Simulated Skewness (Standard Error)
Design in Table 3.5				
τ_1	0.98950(0.00240)	0.00320	0.00574	0.14233(0.07734)
τ_2	−0.99129(0.00240)	0.00322	0.00578	−0.18005(0.07734)
τ_{12}	−0.02435(0.00233)	0.00321	0.00543	0.17092(0.07734)
ν	0.49902(0.00168)	0.00702	0.00282	0.31447(0.07734)
Design in Table 3.6				
τ_1	0.99104(0.00482)	0.00549	0.02319	0.60368(0.07734)
τ_2	−0.96591(0.00446)	0.00729	0.01985	−0.28153(0.07734)
τ_{12}	−0.03756(0.00438)	0.00729	0.01919	0.30541(0.07734)
ν	0.47858(0.00231)	0.01408	0.00533	0.45193(0.07734)
Design in Table 3.11				
τ_1	0.98549(0.00141)	0.00107	0.00198	0.08501(0.07734)
τ_2	−0.98395(0.00134)	0.00107	0.00180	−0.02499(0.07734)
τ_{12}	−0.02720(0.00133)	0.00107	0.00176	0.05257(0.07734)
ν	0.49884(0.00093)	0.00234	0.00086	0.05704(0.07734)
Design in Table 3.12				
τ_1	0.98751(0.00282)	0.00400	0.00793	0.18660(0.07734)
τ_2	−0.98777(0.00272)	0.00411	0.00738	−0.10613(0.07734)
τ_{12}	−0.03117(0.00257)	0.00404	0.00661	0.17683(0.07734)
ν	0.49696(0.00179)	0.00936	0.00322	0.12208(0.07734)

Table 3.13: Summary statistics for $\tau_1 = 1, \tau_2 = -1, \tau_{12} = 0$, and $\nu = 0.5$.

Parameter	Simulated Mean (Standard Error)	Theoretical Variance	Simulated Variance	Simulated Skewness (Standard Error)
Design in Table 3.5				
τ_1	0.99699(0.00236)	0.00329	0.00559	0.10950(0.07734)
τ_2	−0.99738(0.00239)	0.00331	0.00574	−0.22312(0.07734)
τ_{12}	−0.27001(0.00233)	0.00330	0.00541	0.10207(0.07734)
ν	0.50699(0.00175)	0.00702	0.00307	0.31468(0.07734)
Design in Table 3.6				
τ_1	1.01073(0.00513)	0.00575	0.02628	0.57665(0.07734)
τ_2	−0.98728(0.00462)	0.00764	0.02132	−0.31215(0.07734)
τ_{12}	−0.28134(0.00442)	0.00764	0.01950	0.29241(0.07734)
ν	0.49149(0.00271)	0.01407	0.00732	0.60662(0.07734)
Design in Table 3.11				
τ_1	0.99521(0.00137)	0.00110	0.00189	0.13158(0.07734)
τ_2	−0.99530(0.00138)	0.00110	0.00191	−0.07656(0.07734)
τ_{12}	−0.27166(0.00133)	0.00110	0.00176	0.08612(0.07734)
ν	0.50342(0.00098)	0.00234	0.00095	0.14545(0.07734)
Design in Table 3.12				
τ_1	1.00344(0.00302)	0.00418	0.00912	0.23760(0.07734)
τ_2	−0.99695(0.00260)	0.00431	0.00677	−0.11871(0.07734)
τ_{12}	−0.27676(0.00253)	0.00422	0.00640	0.00381(0.07734)
ν	0.50901(0.00207)	0.00936	0.00429	0.42764(0.07734)

Table 3.14: Summary statistics for $\tau_1 = 1, \tau_2 = -1, \tau_{12} = -0.25$, and $\nu = 0.5$.

3.A Proof that $\mathbf{j_L^T z} = 0$ for the generalised Davidson Ties Model

We begin by recalling that

$$z_i = \sum_{i \in C} \left(w_{i|C} + \sum_{\{i,j_2,\ldots,j_x\} \subseteq C} \frac{1}{x} w_{0(i,j_2,\ldots,j_x)|C} - sn_C \widehat{\pi}_i \frac{\widehat{\partial D_C}}{\partial \pi_i} \right),$$

where

$$\frac{\widehat{\partial D_C}}{\partial \pi_i} = \frac{sn_C}{D_C} \left(1 + \sum_{\{i,j_2,\ldots,j_x\} \subseteq C} \frac{\nu \widehat{\pi_{j_2}} \times \ldots \times \widehat{\pi_{j_x}}}{\sqrt[x]{\widehat{\pi_i}\widehat{\pi_{j_2}} \ldots \widehat{\pi_{j_x}}}} \right)$$

for $2 \leq x \leq m$. Now, the vector $\boldsymbol{z}$ contains the values for z_i for each possible item T_i. Then

$$\begin{aligned}
\boldsymbol{j}_L^T \boldsymbol{z} &= \sum_{i=1}^{t} z_i \\
&= \sum_C \left(\sum_j w_{j|C} + \sum_{\{j_1,j_2,\ldots,j_x\} \subseteq C} w_{0(j_1,j_2,\ldots,j_x)|C} \right) \\
&\quad - \sum_C \frac{sn_C}{D_C} \left(\sum_j \widehat{\pi}_j + \sum_{\{j_1,j_2,\ldots,j_x\} \subseteq C} \frac{\widehat{\nu}\widehat{\pi_{j_1}} \ldots \widehat{\pi_{j_x}}}{\sqrt[x]{(\widehat{\pi_{j_1}} \ldots \widehat{\pi_{j_x}})^{x-1}}} \right) \\
&= \sum_C sn_C - \sum_C \frac{sn_C}{D_C} \left(\sum_j \widehat{\pi}_j + \sum_{\{j_1,j_2,\ldots,j_x\} \subseteq C} \widehat{\nu} \sqrt[x]{\widehat{\pi_{j_1}} \ldots \widehat{\pi_{j_x}}} \right) \\
&= \sum_C sn_C - \sum_C \frac{sn_C}{D_C} D_C \\
&= \sum_C sn_C - \sum_C sn_C \\
&= 0,
\end{aligned}$$

as required.

3.B Proof that the generalised Davidson Ties Model does not violate El-Helbawy and Bradley [1978] Conditions

In order to apply the results relating to associated populations, we need to show that $C(\boldsymbol{\pi}_0, \nu)$ is positive definite, as El-Helbawy and Bradley [1978] did.

■ **THEOREM 3.B.1.**
The C matrix for the estimation of a set of contrasts $B_h \boldsymbol{\gamma}$ and ν, where

$$B = \begin{bmatrix} B_h & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix}$$

is positive definite.

Proof. McFadden [1973] show that $C(\boldsymbol{\pi}_0)_{\text{MNL}}$ is positive definite. Then the eigenvalues of $C(\boldsymbol{\pi}_0)_{\text{MNL}}$, $\lambda_1, \ldots, \lambda_p$ are all positive. If we consider the matrix

$$C(\boldsymbol{\pi}_0, \nu)_{\text{DAV}} = \begin{bmatrix} Q(m,\nu) C(\boldsymbol{\pi}_0)_{\text{MNL}} & \mathbf{0} \\ \mathbf{0} & \Lambda_{\nu\nu}(m,\nu) \end{bmatrix},$$

where

$$\Lambda_{\nu\nu}(m,\nu) = \frac{m\sum_{x=2}^{m}\binom{m}{x}}{\nu(m+\nu\sum_{x=2}^{m}\binom{m}{x})^2} > 0$$

Then

$$\det(C(\boldsymbol{\pi}_0,\nu)_{\text{DAV}} - I_{p+1}) = Q(m,\nu)^p \det(C(\boldsymbol{\pi}_0)_{\text{B–T}} - I_p) \times (\Lambda_{\nu\nu}(m,\nu) - 1),$$

will have roots at $\lambda_1, \ldots, \lambda_p, \Lambda_{\nu\nu}(m,\nu)$. We already know that $\lambda_1, \ldots, \lambda_p$ are positive, and $\Lambda_{\nu\nu}(m,\nu)$ is positive for all $\nu > 0$. Since we assume that $\nu > 0$ anyway, all of the eigenvalues are positive. Therefore the C matrix for the generalised Davidson ties model is positive definite. □

Chapter 4

Choice Models that Incorporate Position Effects

The idea that the order of presentation of information can influence responses is well established in questionnaire design. A good discussion of this influence is given in Kalton et al. [1978]. The authors suggest that options presented earlier in a set of alternatives will be selected more often than those appearing later in the set, all other things remaining equal. This is reminiscent of the donkey vote in elections. These ideas also appear in the design of tournaments where the home team is expected to have an advantage.

A choice experiment is similar to a questionnaire in this regard since in a choice experiment we present a set of m alternatives to choose from on each occasion. Given this similarity, it may be useful to incorporate position effects into the choice model. So far, we have assumed the position that an item occupies within the choice set is immaterial. This chapter develops optimal design theory for the Davidson–Beaver position effects model, a model which incorporates the effect of the position of an item within a choice set.

We start by deriving an expression for the determinant of the information matrix when the Davidson–Beaver position effects model is used, where $m = 2$. We then show that, under a mild restriction, the designs that are optimal for the estimation of a set of attribute effects when the Bradley–Terry model is used are optimal for the estimation of the same set of attribute effects and the position main effect over the same set of competing designs when the Davidson–Beaver position effects model is used.

Throughout this chapter, we continue to use the experiment introduced in Example 2.0.12 to illustrate the results.

4.1 Review of the Davidson–Beaver Position Effects Model

We begin by reviewing the results of Davidson and Beaver [1977]. This section will recap some of the properties of the model that have already been developed in the literature, such as the distribution of the responses, the maximum likelihood estimates, and the information matrix for the estimation of this model. The methods presented here will be used when the model is generalised in Chapter 5. We will also use these results to show that the optimal designs for the estimation of the Bradley–Terry model, as presented in Street et al. [2001] and Burgess and

Street [2003], are also optimal when the Davidson–Beaver position effects model is used.

Recall from Section 1.1 that Davidson and Beaver [1977] proposed a model that extended the Bradley–Terry model to incorporate position effects. Position effects are incorporated by introducing two additional parameters ψ_1 and ψ_2. When multiplying the merit of an item, π_i, these parameters measure the effect of an item being presented in the first position of the choice set and in the second position of the choice set respectively. While the model in Davidson and Beaver [1977] scaled $\psi_1 \to 1$ and $\psi_2 \to \psi_2/\psi_1$ without loss of generality, we will not make this restriction here. Instead, we estimate orthogonal polynomial contrasts of the position effects, which will be easier to generalise to an arbitrary choice set size.

We have the following probabilities associated with the possible decisions when the ordered choice set $C = (T_{i_1}, T_{i_2})$ is used.

$$P(T_{i_1}|C) = \frac{\psi_1 \pi_{i_1}}{\psi_1 \pi_{i_1} + \psi_2 \pi_{i_2}}$$
$$P(T_{i_2}|C) = \frac{\psi_2 \pi_{i_2}}{\psi_1 \pi_{i_1} + \psi_2 \pi_{i_2}}$$

We will now see how these probabilities apply to our example.

■ **EXAMPLE 4.1.1.**

There are 12 possible ordered choice sets from the ordered pairs of items listed in Example 2.0.12. If we consider the ordered choice set $(00, 11)$, for instance, then the probability of choosing item 00 is

$$P(00|(00,11)) = \frac{\psi_1 \pi_{00}}{\psi_1 \pi_{00} + \psi_2 \pi_{11}},$$

and the probability of choosing item 11 is

$$P(11|(00,11)) = \frac{\psi_2 \pi_{11}}{\psi_1 \pi_{00} + \psi_2 \pi_{11}}.$$

□

In their 1977 paper, Davidson and Beaver derived the log–likelihood function and information matrix for their position effects model. For the benefit of the reader, we provide a detailed derivation here. We will use this method to derive the information matrix for the generalised Davidson–Beaver position effects model in Chapter 5.

Suppose that there are t items in total and that these are shown to the respondent in pairs. In each choice set, we ask the respondent to choose the item that they prefer. We define indicator variables $w_{i_a|C,\alpha}$ for subject α and ordered choice set $C = (T_{i_1}, T_{i_2})$ to represent the respondent's choice. Thus we let

$$w_{i_1|C,\alpha} = \begin{cases} 1 & \text{if respondent } \alpha \text{ selects item } T_{i_1} \text{ from the ordered choice set } C, \\ 0 & \text{otherwise,} \end{cases}$$
$$w_{i_2|C,\alpha} = \begin{cases} 1 & \text{if respondent } \alpha \text{ selects item } T_{i_2} \text{ from the ordered choice set } C, \\ 0 & \text{otherwise,} \end{cases}$$

where $w_{i_1|C,\alpha} + w_{i_2|C,\alpha} = 1$, since there are no repeated choice sets and we do not have an opt–out process. For respondent α, the probability density function for their response to the ordered choice set $C = (T_{i_1}, T_{i_2})$ is

$$f_{C,\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}) = \frac{(\psi_1 \pi_{i_1})^{w_{i_1|C,\alpha}} \times (\psi_2 \pi_{i_2})^{w_{i_2|C,\alpha}}}{(\psi_1 \pi_{i_1} + \psi_2 \pi_{i_2})^{n_C}},$$

where $\boldsymbol{w} = [w_{i_1|C,\alpha}, w_{i_2|C,\alpha}]$, $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_t)^T$, $\boldsymbol{\psi} = (\psi_1, \psi_2)^T$, and n_C is an indicator that equals 1 if the ordered choice set C appears in the experiment and 0 if it does not. We assume that n_C are the same for all respondents. For consistency, if the ordered choice set C does not appear in the experiment then we define

$$w_{i_1|C,\alpha} = w_{i_2|C,\alpha} = 0.$$

The derivatives of $\ln(f_{C,\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}))$ with respect to each of the parameters are

$$\begin{aligned}
\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}))}{\partial \pi_{i_1}} &= \frac{w_{i_1|C,\alpha}}{\pi_{i_1}} + \frac{n_C \psi_1}{(\psi_1 \pi_{i_1} + \psi_2 \pi_{i_2})}, \\
\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}))}{\partial \pi_{i_2}} &= \frac{w_{i_2|C,\alpha}}{\pi_{i_2}} + \frac{n_C \psi_2}{(\psi_1 \pi_{i_1} + \psi_2 \pi_{i_2})}, \\
\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}))}{\partial \pi_{i_j}} &= 0 \text{ for } T_{i_j} \notin C, \\
\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}))}{\partial \psi_1} &= \frac{w_{i_1|C,\alpha}}{\psi_1} + \frac{n_C \pi_{i_1}}{(\psi_1 \pi_{i_1} + \psi_2 \pi_{i_2})},
\end{aligned}$$

and

$$\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}))}{\partial \psi_2} = \frac{w_{i_2|C,\alpha}}{\psi_2} + \frac{n_C \pi_{i_2}}{(\psi_1 \pi_{i_1} + \psi_2 \pi_{i_2})}.$$

We will use these derivatives later to derive the entries of the information matrix for the estimation of this model. We now turn our attention to the MLEs for this model.

Since the likelihood function is the product of the density functions over all distinct choice sets and over all respondents, we have

$$\begin{aligned}
L(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}) &= \prod_{\alpha=1}^{s} \prod_{C} f_{C,\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}) \\
&= \prod_{C} \frac{(\psi_1 \pi_{i_1})^{w_{i_1|C}} (\psi_2 \pi_{i_2})^{w_{i_2|C}}}{(\psi_1 \pi_{i_1} + \psi_2 \pi_{i_2})^{s n_C}},
\end{aligned}$$

where the product is over the the set of distinct choice sets, and $w_{i_1|C} = \sum_{\alpha=1}^{s} w_{i_1|C,\alpha}$ and $w_{i_2|C} = \sum_{\alpha=1}^{s} w_{i_2|C,\alpha}$.

To maximise the likelihood function, we need to set up a Lagrangian function to incorporate the restrictions placed on this model. For the purposes of convergence, we impose the constraint present in the Bradley–Terry model. Thus

$$\sum_{i=1}^{t} \ln(\pi_i) = 0.$$

Similarly, we place the constraint

$$\ln(\psi_1) + \ln(\psi_2) = 0$$

on the position effects to ensure convergence. We also constrain the contrasts that are assumed to be negligible to equal 0. If we let B_a be the matrix containing the coefficients of these contrasts, then we have

$$B_a \boldsymbol{\gamma} = 0,$$

where $\boldsymbol{\gamma}$ is a vector containing $\gamma_i = \ln(\pi_i)$ for $i = 1, 2, \ldots, t$. This gives the Lagrangian

$$\begin{aligned}
G(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}) &= \sum_{C} \big[w_{i_1|C}(\ln(\pi_{i_1}) + \ln(\psi_1)) + w_{i_2|C}(\ln(\pi_{i_2}) + \ln(\psi_2)) \\
&\quad - s n_C \ln(\psi_1 \pi_{i_1} + \psi_1 \pi_{i_1})\big] + \kappa_1 \sum_{i=1}^{t} \gamma_i + \kappa_2 \sum_{a=1}^{m} \ln(\psi_a) + [\kappa_3 \ldots \kappa_{h+2}] B_a \boldsymbol{\gamma},
\end{aligned}$$

where there are h contrasts in B_a, and $\kappa_1, \ldots, \kappa_{h+2}$ are Lagrange multipliers. If we differentiate $G(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi})$ with respect to π_i, we obtain

$$\frac{\partial G(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi})}{\partial \pi_i} = \sum_{C|T_i \in C} \left\{ \frac{w_{i|C}}{\pi_i} - \frac{sn_C \psi_{a_i}}{(\psi_1 \pi_{i_1} + \psi_2 \pi_{i_2})} \right\} + \frac{\kappa_1}{\pi_i} + \sum_{x=1}^{h} \frac{\kappa_{x+2}(B_a)_{xi}}{\pi_i},$$

where ψ_{a_i} is the position effect corresponding to the position of item T_i in choice set C, and π_{i_a} is the merit of the item in the a^{th} position of choice set C. If we differentiate $G(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi})$ with respect to ψ_a, we obtain

$$\frac{\partial G(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi})}{\partial \psi_a} = \sum_{C} \left\{ \frac{w_{i_a|C}}{\psi_a} - \frac{sn_C \pi_{i_a}}{(\psi_1 \pi_{i_1} + \psi_2 \pi_{i_2})} \right\} + \frac{\kappa_2}{\psi_a}.$$

We obtain the MLEs by setting these derivatives equal to zero and solving simultaneously. We can simplify this problem by using matrix notation. Suppose that we let

$$z_i = \sum_{C|T_i \in C} w_{i|C} - \frac{sn_C \hat{\psi}_{a_i} \hat{\pi}_i}{\hat{\psi}_1 \hat{\pi}_{i_1} + \hat{\psi}_2 \hat{\pi}_{i_2}}.$$

Then, by multiplying Equation 4.1 by $\hat{\pi}_{i_j}$, we obtain

$$z_i + \kappa_1 + \sum_{x=1}^{h} \kappa_{x+2}(B_a)_{xi} = 0.$$

This gives the system

$$\boldsymbol{z} + \kappa_1 \boldsymbol{j}_L^T + B_a^T \boldsymbol{\kappa} = \mathbf{0}_L, \tag{4.1}$$

as a subset of the normal equations, where $\boldsymbol{z} = (z_1, \ldots, z_t)^T$ and $\boldsymbol{\kappa} = (\kappa_3, \ldots, \kappa_{h+2})^T$. Similarly, if we let

$$p_a = \sum_{C} w_{i_a|C} - \frac{sn_C \hat{\psi}_a \hat{\pi}_{i_a}}{(\hat{\psi}_1 \hat{\pi}_{i_1} + \hat{\psi}_2 \hat{\pi}_{i_2})},$$

then we obtain

$$\boldsymbol{p} + \kappa_2 \boldsymbol{j}_m^T = \mathbf{0}_m$$

as the remaining normal equations, where $\boldsymbol{p} = (p_1, p_2)^T$. Appendix 4.A proves that $\boldsymbol{j}_L^T \boldsymbol{z} = 0$, and it is obvious that $\boldsymbol{j}_L^T B_a^T = \mathbf{0}$ since the rows of B_a are contrast coefficients. It follows that $\kappa_1 = 0$.

We pre–multiply Equation 4.1 by B_a to obtain

$$\boldsymbol{\kappa} = -B_a^T.$$

Substituting this back into Equation 4.1, we get

$$(I - B_a^T B_a)\boldsymbol{z} = \mathbf{0}_L,$$

and

$$\boldsymbol{p} + \kappa_2 \boldsymbol{j}_m^T = \mathbf{0}_m$$

as the normal equations. These can be solved simultaneously to find the MLEs.

We now look at how this can be applied to our example to obtain maximum likelihood estimates.

■ **EXAMPLE 4.1.2.**

Recall the experiment presented in Example 2.0.12. Suppose that we present the ordered choice sets given in Table 4.1 to 50 respondents. The final two columns of Table 4.1 gives a possible set of summarised responses for this experiment and the corresponding likelihood function is

$$L(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}) = \frac{(\psi_1\pi_{00})^{38}(\psi_2\pi_{11})^{12}}{(\psi_1\pi_{00}+\psi_2\pi_{11})^{50}} \times \frac{(\psi_1\pi_{01})^{11}(\psi_2\pi_{10})^{39}}{(\psi_1\pi_{01}+\psi_2\pi_{10})^{50}} \times \frac{(\psi_1\pi_{10})^{46}(\psi_2\pi_{01})^{4}}{(\psi_1\pi_{10}+\psi_2\pi_{01})^{50}} \times \frac{(\psi_1\pi_{11})^{34}(\psi_2\pi_{00})^{16}}{(\psi_1\pi_{11}+\psi_2\pi_{00})^{50}}.$$

Now suppose that we are interested in the estimation of main effects of the attributes and the position main effect only, then

$$B_a = \frac{1}{2}[1 \quad -1 \quad -1 \quad 1],$$

and we have the constraints

$$\begin{aligned}
\ln(\pi_{00}) + \ln(\pi_{01}) + \ln(\pi_{10}) + \ln(\pi_{11}) &= 0,\\
\ln(\pi_{00}) - \ln(\pi_{01}) - \ln(\pi_{10}) + \ln(\pi_{11}) &= 0,\\
\ln(\psi_1) + \ln(\psi_2) &= 0.
\end{aligned}$$

This gives the Lagrangian

$$\begin{aligned}
G(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}) = {} & 38(\ln(\pi_{00}) + \ln(\psi_1)) + 12(\ln(\pi_{11}) + \ln(\psi_2)) - 50\ln((\psi_1\pi_{00} + \psi_2\pi_{11})\\
& +11(\ln(\pi_{01}) + \ln(\psi_1)) + 39(\ln(\pi_{10}) + \ln(\psi_2)) - 50\ln(\psi_1\pi_{01} + \psi_2\pi_{10})\\
& +46(\ln(\pi_{10}) + \ln(\psi_1)) + 4(\ln(\pi_{01}) + \ln(\psi_2)) - 50\ln((\psi_1\pi_{10} + \psi_2\pi_{01})\\
& +34(\ln(\pi_{11}) + \ln(\psi_1)) + 16(\ln(\pi_{00}) + \ln(\psi_2)) - 50\ln(\psi_1\pi_{11} + \psi_2\pi_{00})\\
& +\kappa_1(\ln(\pi_{00}) + \ln(\pi_{01}) + \ln(\pi_{10}) + \ln(\pi_{11})) + \kappa_2(\ln(\psi_1) + \ln(\psi_2))\\
& +\kappa_3(\ln(\pi_{00}) - \ln(\pi_{01}) - \ln(\pi_{10}) + \ln(\pi_{11})).
\end{aligned}$$

We differentiate $G(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi})$ with respect to each π_i to give

$$\begin{aligned}
\frac{\partial G(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi})}{\partial \pi_{00}} &= \frac{38+16}{\pi_{00}} - \frac{50\psi_1}{\psi_1\pi_{00}+\psi_2\pi_{11}} - \frac{50\psi_2}{\psi_1\pi_{11}+\psi_2\pi_{00}} + \frac{\kappa_1}{\pi_{00}} + \frac{\kappa_3}{\pi_{00}},\\
\frac{\partial G(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi})}{\partial \pi_{01}} &= \frac{11+4}{\pi_{01}} - \frac{50\psi_1}{\psi_1\pi_{01}+\psi_2\pi_{10}} - \frac{50\psi_2}{\psi_1\pi_{10}+\psi_2\pi_{01}} + \frac{\kappa_1}{\pi_{01}} - \frac{\kappa_3}{\pi_{01}},\\
\frac{\partial G(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi})}{\partial \pi_{10}} &= \frac{39+46}{\pi_{10}} - \frac{50\psi_1}{\psi_1\pi_{01}+\psi_2\pi_{10}} - \frac{50\psi_2}{\psi_1\pi_{10}+\psi_2\pi_{01}} + \frac{\kappa_1}{\pi_{10}} - \frac{\kappa_3}{\pi_{10}},
\end{aligned}$$

and

$$\frac{\partial G(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi})}{\partial \pi_{11}} = \frac{12+34}{\pi_{11}} - \frac{50\psi_1}{\psi_1\pi_{11}+\psi_2\pi_{00}} - \frac{50\psi_2}{\psi_1\pi_{00}+\psi_2\pi_{11}} + \frac{\kappa_1}{\pi_{11}} + \frac{\kappa_3}{\pi_{11}}.$$

Option 1	Option 2	$\mathbf{T_1}$	$\mathbf{T_2}$
0 0	1 1	$w_{00\|C} = 38$	$w_{11\|C} = 12$
0 1	1 0	$w_{01\|C} = 11$	$w_{10\|C} = 39$
1 0	0 1	$w_{01\|C} = 46$	$w_{10\|C} = 4$
1 1	0 0	$w_{00\|C} = 34$	$w_{11\|C} = 16$

Table 4.1: A set of responses for the experiment in Example 4.1.2.

Differentiating $G(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi})$ with respect to each ψ_i gives

$$\begin{aligned}\frac{\partial G(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi})}{\partial \psi_1} &= \frac{38+11+46+34}{\psi_1} - \frac{50\pi_{00}}{\psi_1\pi_{00}+\psi_2\pi_{11}} - \frac{50\pi_{01}}{\psi_1\pi_{01}+\psi_2\pi_{10}} \\ &\quad - \frac{50\pi_{10}}{\psi_1\pi_{10}+\psi_2\pi_{01}} - \frac{50\pi_{11}}{\psi_1\pi_{11}+\psi_2\pi_{00}} + \frac{\kappa_2}{\psi_1},\end{aligned}$$

and

$$\begin{aligned}\frac{\partial G(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi})}{\partial \psi_2} &= \frac{12+39+4+16}{\psi_2} - \frac{50\pi_{11}}{\psi_1\pi_{00}+\psi_2\pi_{11}} - \frac{50\pi_{10}}{\psi_1\pi_{01}+\psi_2\pi_{10}} \\ &\quad - \frac{50\pi_{01}}{\psi_1\pi_{10}+\psi_2\pi_{01}} - \frac{50\pi_{00}}{\psi_1\pi_{11}+\psi_2\pi_{00}} + \frac{\kappa_2}{\psi_2}.\end{aligned}$$

If we set each of these to 0 and solve iteratively we obtain the MLEs for the entries in $\boldsymbol{\pi}$ and $\boldsymbol{\psi}$. If we let τ_1 be the main effect of the first attribute and τ_2 the main effect of the second attribute, and ψ_{L} be the main effect of position, then we find

$$\widehat{\tau_1} = 0.446 \qquad \widehat{\tau_2} = -0.541 \qquad \widehat{\psi_{\text{L}}} = -0.419. \qquad \square$$

4.2 Properties of the Davidson–Beaver Position Effects Model

In this section, we complete the construction of the information matrix for the estimation of the entries in $\boldsymbol{\pi}$ and $\boldsymbol{\psi}$. We begin by deriving expressions for the expectations, variances and covariances of the selection indicators $\boldsymbol{w}$. We then use these expressions to simplify the information matrix for the estimation of the entries in $\boldsymbol{\pi}$ and $\boldsymbol{\psi}$.

Recall that $w_{i_1|C,\alpha}$ and $w_{i_2|C,\alpha}$ are the selection indicators for the choice made by respondent α when presented with the ordered choice set $C = (T_{i_1}, T_{i_2})$. These selection indicators each have a Bernoulli distribution with expectations

$$\mathcal{E}_\pi(w_{i_1|C,\alpha}) = \frac{\psi_1\pi_{i_1}}{\psi_1\pi_{i_1}+\psi_2\pi_{i_2}},$$

and

$$\mathcal{E}_\pi(w_{i_2|C,\alpha}) = \frac{\psi_2\pi_{i_2}}{\psi_1\pi_{i_1}+\psi_2\pi_{i_2}} \tag{4.2}$$

respectively. The variances of these selection indicators are

$$\text{Var}_\pi(w_{i_1|C,\alpha}) = \frac{\psi_1\psi_2\pi_{i_1}\pi_{i_2}}{(\psi_1\pi_{i_1}+\psi_2\pi_{i_2})^2},$$

and

$$\text{Var}_\pi(w_{i_2|C,\alpha}) = \frac{\psi_1\psi_2\pi_{i_1}\pi_{i_2}}{(\psi_1\pi_{i_1}+\psi_2\pi_{i_2})^2}. \tag{4.3}$$

Next we derive the covariances between the selection indicators. Consider the covariance of the two selection indicators for the selection of item T_{i_1} from the ordered choice set $C = (T_{i_1}, T_{i_3})$, and the item T_{i_2} from the ordered choice set $C' = (T_{i_4}, T_{i_2})$. If the selections made in two distinct choice sets are independent, then

$$\begin{aligned}\text{Cov}_\pi(w_{i_1|C,\alpha}, w_{i_2|C',\alpha}) &= \mathcal{E}_\pi\Big(\big(w_{i_1|C,\alpha} - \mathcal{E}_\pi(w_{i_1|C,\alpha})\big)\big(w_{i_2|C',\alpha} - \mathcal{E}_\pi(w_{i_2|C',\alpha})\big)\Big) \\ &= \begin{cases} \mathcal{E}_\pi\Big(\big(w_{i_1|C,\alpha} - \mathcal{E}_\pi(w_{i_1|C,\alpha})\big)\big(w_{i_2|C,\alpha} - \mathcal{E}_\pi(w_{i_2|C,\alpha})\big)\Big), & \text{if } C = C'; \\ 0, & \text{otherwise.}\end{cases}\end{aligned}$$

We expand this expectation and notice that only one outcome is possible. Therefore we have $\mathcal{E}_\pi(w_{i_1|i_1 i_2\alpha} w_{i_2|i_1 i_2\alpha}) = 0$ once again, and hence

$$\text{Cov}_\pi(w_{i_1|C,\alpha}, w_{i_2|C',\alpha}) = \begin{cases} \frac{-\psi_1\psi_2\pi_{i_1}\pi_{i_2}}{(\psi_1\pi_{i_1}+\psi_2\pi_{i_2})^2}, & \text{if } C = C' \text{ and } i_1 \neq i_2, \\ \text{Var}_\pi(w_{i_1|C,\alpha}), & \text{if } C = C' \text{ and } i_1 = i_2, \\ 0, & \text{otherwise.} \end{cases} \quad (4.4)$$

We now find the expectations, variances and covariances for the selection indicators in our example.

■ **EXAMPLE 4.2.1.**

Consider the experiment in Example 4.1.2. In particular, consider the first choice set, $C = (00, 11)$. The expected values for the selection indicators for each choice are

$$\mathcal{E}_\pi(w_{00|C,\alpha}) = \frac{\psi_1\pi_{00}}{\psi_1\pi_{00} + \psi_2\pi_{11}},$$

and

$$\mathcal{E}_\pi(w_{11|C,\alpha}) = \frac{\psi_2\pi_{11}}{\psi_1\pi_{00} + \psi_2\pi_{11}}.$$

The variances of the selection indicators for each choice are

$$\text{Var}_\pi(w_{00|C,\alpha}) = \frac{\psi_1\psi_2\pi_{00}\pi_{11}}{(\psi_1\pi_{00} + \psi_2\pi_{11})^2},$$

and

$$\text{Var}_\pi(w_{11|C,\alpha}) = \frac{\psi_1\psi_2\pi_{00}\pi_{11}}{(\psi_1\pi_{00} + \psi_2\pi_{11})^2}.$$

The covariance of these selection indicators is

$$\text{Cov}_\pi(w_{00|C,\alpha}, w_{11|C,\alpha}) = \frac{-\psi_1\psi_2\pi_{00}\pi_{11}}{(\psi_1\pi_{00} + \psi_2\pi_{11})^2}.$$

□

Next we construct the information matrix for the Davidson–Beaver position effects model. This construction is easier if we partition the information matrix into four blocks. $I_{\pi\pi}(\boldsymbol{\pi}, \boldsymbol{\psi})$ is a $t \times t$ matrix that contains minus the expected value of the second derivatives of the density function with respect to two of the entries in $\boldsymbol{\pi}$. $I_{\psi\psi}(\boldsymbol{\pi}, \boldsymbol{\psi})$ is a 2×2 matrix that contains minus the expected value of the second derivatives of the density function with respect to two entries in $\boldsymbol{\psi}$. $I_{\pi\psi}(\boldsymbol{\pi}, \boldsymbol{\psi})$ and $I_{\psi\pi}(\boldsymbol{\pi}, \boldsymbol{\psi})$ contain minus the expected value of the second derivatives with respect to one entry in $\boldsymbol{\pi}$ and one entry in $\boldsymbol{\psi}$. The partitioned matrix is

$$I(\boldsymbol{\pi}, \boldsymbol{\psi}) = \begin{bmatrix} I_{\pi\pi}(\boldsymbol{\pi}, \boldsymbol{\psi}) & I_{\pi\psi}(\boldsymbol{\pi}, \boldsymbol{\psi}) \\ I_{\psi\pi}(\boldsymbol{\pi}, \boldsymbol{\psi}) & I_{\psi\psi}(\boldsymbol{\pi}, \boldsymbol{\psi}) \end{bmatrix},$$

where $I_{\pi\psi}(\boldsymbol{\pi}, \boldsymbol{\psi})$ is a $2 \times t$ matrix, and $I_{\psi\pi}(\boldsymbol{\pi}, \boldsymbol{\psi}) = (I_{\pi\psi}(\boldsymbol{\pi}, \boldsymbol{\psi}))^T$.

El-Helbawy and Bradley [1978] state that, under some mild regularity conditions, as given in Section 1.1, the $(i, j)^{\text{th}}$ entry of the information matrix for a discrete choice experiment without position effects is

$$I(\boldsymbol{\pi})_{ij} = \sum_{q=1}^{t-1} \sum_{r=q+1}^{t} \frac{n_{qr}}{N} \mathcal{E}_\pi\Big(\Big(\frac{\partial \ln(f_{qr\alpha}(\boldsymbol{\pi}, \boldsymbol{w}))}{\partial \pi_i}\Big)\Big(\frac{\partial \ln(f_{qr\alpha}(\boldsymbol{\pi}, \boldsymbol{w}))}{\partial \pi_j}\Big)\Big).$$

We now use this expression, and the results given in Equations 4.2, 4.3, and 4.4, to evaluate some generic cells in each block matrix of $I(\boldsymbol{\pi}, \boldsymbol{\psi})$. To assist the generalisation in Chapter 5 we take the sum over all choice sets rather than the pairs of items, and modify the notation for n_{qr} and $f_{qr\alpha}(\boldsymbol{\pi}, \boldsymbol{w})$ accordingly.

We begin with the $I_{\pi\pi}(\boldsymbol{\pi}, \boldsymbol{\psi})$. In this block matrix we need to consider the diagonal and off–diagonal entries separately. We begin with the generic off–diagonal entry. Consider

$$I_{\pi\pi}(\boldsymbol{\pi}, \boldsymbol{\psi})_{ij} = \sum_C \frac{n_C}{N} \mathcal{E}_\pi\Big(\Big(\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}))}{\partial \pi_i}\Big)\Big(\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}))}{\partial \pi_j}\Big)\Big).$$

We recall that the derivative of the density function is zero if we differentiate with respect to a π_i that is associated with an item that is not in the choice set. Then, unless both items T_i and T_j appear in the choice set, the product of the derivatives will be equal to zero. Using this, we have

$$\begin{aligned} I_{\pi\pi}(\boldsymbol{\pi}, \boldsymbol{\psi})_{ij} &= \frac{n_{(i,j)}}{N} \mathcal{E}_\pi\Big(\Big(\frac{\partial \ln(f_{(i,j),\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}))}{\partial \pi_i}\Big)\Big(\frac{\partial \ln(f_{(i,j),\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}))}{\partial \pi_j}\Big)\Big) \\ &\quad + \frac{n_{(j,i)}}{N} \mathcal{E}_\pi\Big(\Big(\frac{\partial \ln(f_{(j,i),\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}))}{\partial \pi_i}\Big)\Big(\frac{\partial \ln(f_{(j,i),\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}))}{\partial \pi_j}\Big)\Big). \end{aligned}$$

We observe that

$$\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}))}{\partial \pi_i} = \frac{w_{i|C,\alpha}}{\pi_i} - \mathcal{E}_\pi\Big(\frac{w_{i|C,\alpha}}{\pi_i}\Big). \tag{4.5}$$

Then it follows that

$$I_{\pi\pi}(\boldsymbol{\pi}, \boldsymbol{\psi})_{ij} = \frac{n_{(i,j)}}{N} \mathrm{Cov}_\pi\Big(\frac{w_{i|(i,j),\alpha}}{\pi_i}, \frac{w_{j|(i,j),\alpha}}{\pi_j}\Big) + \frac{n_{(j,i)}}{N} \mathrm{Cov}_\pi\Big(\frac{w_{j|(j,i),\alpha}}{\pi_j}, \frac{w_{i|(j,i),\alpha}}{\pi_i}\Big).$$

We now substitute the covariance given by Equation 4.4, giving

$$I_{\pi\pi}(\boldsymbol{\pi}, \boldsymbol{\psi})_{ij} = \frac{-n_{(i,j)}\psi_1\psi_2}{N(\psi_1\pi_i + \psi_2\pi_j)^2} + \frac{-n_{(j,i)}\psi_1\psi_2}{N(\psi_1\pi_j + \psi_2\pi_i)^2}.$$

Next we consider a generic diagonal entry of $I_{\pi\pi}(\boldsymbol{\pi}, \boldsymbol{\psi})$. We have

$$I_{\pi\pi}(\boldsymbol{\pi}, \boldsymbol{\psi})_{ii} = \sum_C \frac{n_C}{N} \mathcal{E}_\pi\Big(\Big(\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}))}{\partial \pi_i}\Big)^2\Big).$$

Again, we observe that the derivative in this expression will be 0 unless the ordered choice set includes the item T_i. So we obtain

$$I_{\pi\pi}(\boldsymbol{\pi}, \boldsymbol{\psi})_{ii} = \sum_{C|T_i \in C} \frac{n_C}{N} \mathcal{E}_\pi\Big(\Big(\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}))}{\partial \pi_i}\Big)^2\Big).$$

Using Equation 4.5, this becomes

$$I_{\pi\pi}(\boldsymbol{\pi}, \boldsymbol{\psi})_{ii} = \sum_{C|T_i \in C} \frac{n_C}{N} \mathrm{Var}_\pi\Big(\frac{w_{i|C\alpha}}{\pi_i}\Big).$$

When we substitute Equation 4.3, we obtain

$$I_{\pi\pi}(\boldsymbol{\pi}, \boldsymbol{\psi})_{ii} = \sum_{C|T_i \in C} \frac{n_C\psi_1\psi_2\pi_{i_1}\pi_{i_2}}{N\pi_i^2(\psi_1\pi_{i_1} + \psi_2\pi_{i_2})^2}.$$

Now we turn our attention to the $I_{\pi\psi}(\boldsymbol{\pi}, \boldsymbol{\psi})$. We begin with

$$I_{\pi\psi}(\boldsymbol{\pi}, \boldsymbol{\psi})_{ia} = \sum_C \frac{n_C}{N} \mathcal{E}_\pi\Big(\Big(\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}))}{\partial \pi_i}\Big)\Big(\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}))}{\partial \psi_a}\Big)\Big),$$

and notice that the product of the derivatives will only be non–zero if the item T_i is in the ordered choice set. Thus we have

$$I_{\pi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})_{ia} = \sum_{C|T_i\in C} \frac{n_C}{N}\mathcal{E}_\pi\Big(\Big(\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}))}{\partial \pi_i}\Big)\Big(\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}))}{\partial \psi_a}\Big)\Big).$$

If we take into account Equation 4.5, and that

$$\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}))}{\partial \psi_a} = \frac{w_{i_a|C,\alpha}}{\psi_a} - \mathcal{E}_\pi\Big(\frac{w_{i_a|C,\alpha}}{\psi_a}\Big), \tag{4.6}$$

where T_{i_a} is the item in the a^{th} position of the ordered choice set, then

$$I_{\pi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})_{ia} = \sum_{C|T_i\in C} \frac{n_C}{N}\text{Cov}_\pi\Big(\frac{w_{i|C,\alpha}}{\pi_i},\frac{w_{i_a|C,\alpha}}{\psi_a}\Big).$$

We substitute in Equations 4.3 and 4.4 to obtain

$$I_{\pi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})_{ia} = \begin{cases} \sum_{i_2\neq i}\frac{n_{(i,i_2)}\psi_2\pi_{i_2}}{N(\psi_1\pi_i+\psi_2\pi_{i_2})^2} - \sum_{i_1\neq i}\frac{n_{(i_1,i)}\psi_2\pi_{i_1}}{N(\psi_1\pi_{i_1}+\psi_2\pi_i)^2}, & \text{if } a=1,\\ \sum_{i_1\neq i}\frac{n_{(i_1,i)}\psi_1\pi_{i_1}}{N(\psi_1\pi_{i_1}+\psi_2\pi_i)^2} - \sum_{i_2\neq i}\frac{n_{(i,i_2)}\psi_1\pi_{i_2}}{N(\psi_1\pi_i+\psi_2\pi_{i_2})^2}, & \text{if } a=2. \end{cases}$$

These expressions give the entries in $I_{\pi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})$. $I_{\psi\pi}(\boldsymbol{\pi},\boldsymbol{\psi})$ is the transpose of $I_{\pi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})$.

Finally, we look at $I_{\psi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})$. Again, it is convenient to consider the generic diagonal and off–diagonal entries separately. We have

$$I_{\psi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})_{1,2} = \sum_{C} \frac{n_C}{N}\mathcal{E}_\pi\Big(\Big(\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}))}{\partial \psi_1}\Big)\Big(\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}))}{\partial \psi_2}\Big)\Big).$$

If we use Equation 4.6, then this simplifies to

$$I_{\psi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})_{1,2} = \sum_{C} \frac{n_C}{N}\text{Cov}_\pi\Big(\frac{w_{i_1|C,\alpha}}{\psi_1},\frac{w_{i_2|C,\alpha}}{\psi_2}\Big).$$

When we substitute in Equation 4.4 and simplify, we obtain

$$I_{\psi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})_{1,2} = \sum_{C} \frac{-n_C\pi_{i_1}\pi_{i_2}}{N(\psi_1\pi_{i_1}+\psi_2\pi_{i_2})^2}.$$

We now look at the diagonal entries of the $I_{\psi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})$ block matrix. We begin with

$$I_{\psi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})_{aa} = \sum_{C} \frac{n_C}{N}\mathcal{E}_\pi\Big(\Big(\frac{\partial \ln(f_{C,\alpha}(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}))}{\partial \psi_a}\Big)^2\Big).$$

If we observe the property in Equation 4.6, then this entry simplifies to

$$I_{\psi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})_{aa} = \sum_{C} \frac{n_C}{N}\text{Var}_\pi\Big(\frac{w_{i_a|C,\alpha}}{\psi_a}\Big).$$

Use Equation 4.3, and simplify, to obtain

$$I_{\psi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})_{aa} = \sum_{C} \frac{n_C\psi_1\psi_2\pi_{i_1}\pi_{i_2}}{N\psi_a^2(\psi_1\pi_{i_1}+\psi_2\pi_{i_2})^2}.$$

Since our ultimate goal is the estimation of main effects and interaction effects, which are linear combinations of the $\boldsymbol{\gamma}=\ln(\boldsymbol{\pi})$, as well as position effects, we first need to construct the information matrix for the estimation of the entries in $\boldsymbol{\gamma}$ and $\boldsymbol{\psi}$. The equivalent information matrix for the estimation of the MNL model was introduced in Section 1.1, and is denoted by $\Lambda(\boldsymbol{\pi})$. For the Davidson–Beaver position effects model we use the same notation, giving

$$P = \begin{bmatrix} \pi_1 & 0 & \dots & 0 & 0 & 0 \\ 0 & \pi_2 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \pi_t & 0 & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 \end{bmatrix},$$

since $\frac{\partial \pi_i}{\partial \gamma_i} = \pi_i$ and $\frac{\partial \psi_a}{\partial \psi_a} = 1$. It is convenient to partition the $\Lambda(\boldsymbol{\pi}, \boldsymbol{\psi})$ matrix in the same way as $I(\boldsymbol{\pi}, \boldsymbol{\psi})$, so

$$\Lambda(\boldsymbol{\pi}, \boldsymbol{\psi}) = \begin{bmatrix} \Lambda_{\gamma\gamma}(\boldsymbol{\pi}, \boldsymbol{\psi}) & \Lambda_{\gamma\psi}(\boldsymbol{\pi}, \boldsymbol{\psi}) \\ \Lambda_{\psi\gamma}(\boldsymbol{\pi}, \boldsymbol{\psi}) & \Lambda_{\psi\psi}(\boldsymbol{\pi}, \boldsymbol{\psi}) \end{bmatrix},$$

where $\Lambda_{\psi\gamma}(\boldsymbol{\pi}, \boldsymbol{\psi}) = (\Lambda_{\gamma\psi}(\boldsymbol{\pi}, \boldsymbol{\psi}))^T$.

If we apply this transformation to each of the generic entries in each block and simplify, we obtain

$$\begin{aligned}
\Lambda_{\gamma\gamma}(\boldsymbol{\pi}, \boldsymbol{\psi})_{ij} &= \frac{-n_{(i,j)}\psi_1\psi_2\pi_i\pi_j}{N(\psi_1\pi_i + \psi_2\pi_j)^2} + \frac{-n_{(j,i)}\psi_1\psi_2\pi_i\pi_j}{N(\psi_1\pi_j + \psi_2\pi_i)^2}, \\
\Lambda_{\gamma\gamma}(\boldsymbol{\pi}, \boldsymbol{\psi})_{ii} &= \sum_{C|T_i \in C} \frac{n_C\psi_1\psi_2\pi_{i_1}\pi_{i_2}}{N(\psi_1\pi_1 + \psi_2\pi_{i_2})^2}, \\
\Lambda_{\gamma\psi}(\boldsymbol{\pi}, \boldsymbol{\psi})_{ia} &= \begin{cases} \sum_{i_2 \neq i} \frac{n_{(i,i_2)}\psi_2\pi_i\pi_{i_2}}{N(\psi_1\pi_i + \psi_2\pi_{i_2})^2} - \sum_{i_1 \neq i} \frac{n_{(i_1,i)}\psi_2\pi_i\pi_{i_1}}{N(\psi_1\pi_{i_1} + \psi_2\pi_i)^2}, & \text{if } a = 1, \\ \sum_{i_1 \neq i} \frac{n_{(i_1,i)}\psi_1\pi_i\pi_{i_1}}{N(\psi_1\pi_{i_1} + \psi_2\pi_i)^2} - \sum_{i_2 \neq i} \frac{n_{(i,i_2)}\psi_1\pi_i\pi_{i_2}}{N(\psi_1\pi_i + \psi_2\pi_{i_2})^2}, & \text{if } a = 2, \end{cases} \\
\Lambda_{\psi\psi}(\boldsymbol{\pi}, \boldsymbol{\psi})_{1,2} &= \sum_C \frac{-n_C\pi_{i_1}\pi_{i_2}}{N(\psi_1\pi_{i_1} + \psi_2\pi_{i_2})^2},
\end{aligned}$$

and

$$\Lambda_{\psi\psi}(\boldsymbol{\pi}, \boldsymbol{\psi})_{aa} = \sum_C \frac{n_C\psi_1\psi_2\pi_{i_1}\pi_{i_2}}{N\psi_a^2(\psi_1\pi_{i_1} + \psi_2\pi_{i_2})^2}.$$

If we make the assumption of equal merits, that is

$$\boldsymbol{\pi} = \boldsymbol{j} = \boldsymbol{\pi}_0,$$

then these expressions simplify further. We leave ψ_1 and ψ_2 unspecified and substitute into the generic entries in each block, to obtain

$$\begin{aligned}
\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{ij} &= \frac{-\psi_1\psi_2}{(\psi_1 + \psi_2)^2} \sum_{C|T_i, T_j \in C} \lambda_C, \\
\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{ii} &= \frac{\psi_1\psi_2}{(\psi_1 + \psi_2)^2} \sum_{C|T_i \in C} \lambda_C, \\
\Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{ia} &= \frac{\psi_b}{(\psi_1 + \psi_2)^2} \sum_{C|T_i \in C} (\lambda_{T_i \text{ in pos } a} - \lambda_{T_i \text{ not in pos } a}), \\
\Lambda_{\psi\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{1,2} &= \frac{-1}{(\psi_1 + \psi_2)^2},
\end{aligned}$$

and

$$\Lambda_{\psi\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{aa} = \frac{\psi_1\psi_2}{\psi_a^2(\psi_1+\psi_2)^2},$$

where $\lambda_C = n_C/N$, $b \neq a$,

$$\lambda_{T_i \text{ in pos } a} = \sum_{C|T_i \text{ in pos } a \text{ of } C} \frac{n_C}{N},$$

and

$$\lambda_{T_i \text{ not in pos } a} = \sum_{C|T_i \in C} \left(\frac{n_C}{N}\right) - \lambda_{T_i \text{ in pos } a}.$$

Now we will construct this information matrix for our example.

■ **EXAMPLE 4.2.2.**

Recall the experiment introduced in Example 2.0.12 and the design introduced in Example 4.1.2. The information matrix for the estimation of the entries in $\boldsymbol{\gamma}$ and $\boldsymbol{\psi}$, under the null hypothesis of equal merits, is given by

$$\Lambda(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = \begin{bmatrix} \frac{\psi_1\psi_2}{2(\psi_1+\psi_2)^2} & 0 & 0 & \frac{-\psi_1\psi_2}{2(\psi_1+\psi_2)^2} & 0 & 0 \\ 0 & \frac{\psi_1\psi_2}{2(\psi_1+\psi_2)^2} & \frac{-\psi_1\psi_2}{2(\psi_1+\psi_2)^2} & 0 & 0 & 0 \\ 0 & \frac{-\psi_1\psi_2}{2(\psi_1+\psi_2)^2} & \frac{\psi_1\psi_2}{2(\psi_1+\psi_2)^2} & 0 & 0 & 0 \\ \frac{-\psi_1\psi_2}{2(\psi_1+\psi_2)^2} & 0 & 0 & \frac{\psi_1\psi_2}{2(\psi_1+\psi_2)^2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{\psi_2}{\psi_1(\psi_1+\psi_2)^2} & \frac{-1}{(\psi_1+\psi_2)^2} \\ 0 & 0 & 0 & 0 & \frac{-1}{(\psi_1+\psi_2)^2} & \frac{\psi_1}{\psi_2(\psi_1+\psi_2)^2} \end{bmatrix}.$$

□

4.3 Representing Options using k Attributes

In this section we consider the construction of the information matrix when contrasts of the entries in $\boldsymbol{\gamma} = \ln(\boldsymbol{\pi})$ and contrasts of the entries in $\boldsymbol{\psi}$ are of interest. In particular, we are interested in contrasts of the entries in $\boldsymbol{\gamma}$ that represent the main effects and interaction effects of the attributes as introduced in Chapter 1.

Ideally, we would like to find the effect of level f_q of attribute q, denoted by β_{q,f_q}, or combinations of attribute levels, on the merit of an item. That is, we want to estimate

$$\boldsymbol{\beta} = (\beta_{1,0}, \beta_{1,1}, \ldots, \beta_{1,\ell_1-1}, \ldots, \beta_{k,\ell_k-1}, \beta_{12,00}, \ldots, \beta_{12\ldots k,\ell_1-1\ldots\ell_k-1}, \psi_1, \ldots, \psi_m)^T.$$

This is not possible however, because $\boldsymbol{\beta}$ is not estimable. It would be better to estimate contrasts of the entries in $\boldsymbol{\beta}$ so that we have a set of estimable contrasts. Suppose that the matrix B contains contrast coefficients that correspond to the coefficients of the effects that are of interest. We can choose the entries in B such that $B\boldsymbol{\beta}$ is estimable.

In general, we are not interested in the interaction between the attribute effects and the position effects, and we assume that contrasts involving entries in both $\boldsymbol{\gamma}$ and $\boldsymbol{\psi}$ will be 0. Thus we can express the matrix of contrast coefficients as the partitioned matrix

$$B = \begin{bmatrix} B_\gamma & \mathbf{0} \\ \mathbf{0} & B_\psi \end{bmatrix},$$

where B_γ contains the contrast coefficients relating to the attribute effects, and B_ψ contains contrast coefficients relating to the position main effect.

The information matrix for the estimation of the contrasts in B is

$$C(\boldsymbol{\pi},\boldsymbol{\psi}) = B\Lambda(\boldsymbol{\pi},\boldsymbol{\psi})B^T,$$

which becomes

$$C(\boldsymbol{\pi}_0,\boldsymbol{\psi}) = \begin{bmatrix} B_\gamma\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0,\boldsymbol{\psi})B_\gamma^T & B_\gamma\Lambda_{\gamma\psi}(\boldsymbol{\pi}_0,\boldsymbol{\psi})B_\psi^T \\ B_\psi\Lambda_{\psi\gamma}(\boldsymbol{\pi}_0,\boldsymbol{\psi})B_\gamma^T & B_\psi\Lambda_{\psi\psi}(\boldsymbol{\pi}_0,\boldsymbol{\psi})B_\psi^T \end{bmatrix}.$$

The terms of the $\Lambda(\boldsymbol{\pi}_0,\boldsymbol{\psi})$ matrix were given in Section 4.2.

Now we find the information matrices in the case where we wish to estimate main effects of the attributes and the position main effect for our example.

■ **EXAMPLE 4.3.1.**

Consider the experiment introduced in Example 2.0.12 and the design introduced in Example 4.1.2 to estimate the main effects of the attributes and the position main effect. The contrast matrix for the estimation of these effects is

$$B = \frac{1}{2}\begin{bmatrix} -1 & -1 & 1 & 1 & 0 & 0 \\ -1 & 1 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\sqrt{2} & \sqrt{2} \end{bmatrix},$$

where B_γ is a 2×4 matrix of contrast coefficients and B_ψ a 1×2 matrix of contrast coefficients. Then the information matrix for the estimation of the main effects and the position main effect is

$$C(\boldsymbol{\pi}_0,\boldsymbol{\psi}) = \begin{bmatrix} \frac{\psi_1\psi_2}{(\psi_1+\psi_2)^2} & 0 & 0 \\ 0 & \frac{\psi_1\psi_2}{(\psi_1+\psi_2)^2} & 0 \\ 0 & 0 & \frac{1}{2\psi_1\psi_2} \end{bmatrix}.$$

Since the information matrix is diagonal, we are able to estimate the main effects of the attributes and the position main effect independently when using this design . □

Now that we have an expression for the information matrix for the estimation of a set of attribute effects that are of interest and the position main effect, we will develop some results on the optimality of designs when using this model.

4.4 Optimal Designs for the Davidson–Beaver Position Effects Model

In this section we compare the information matrix for the estimation of a set of attribute effects when the Bradley–Terry model is used to the information matrix when the Davidson–Beaver position effects model is used. Throughout this section we will assume that the same set of contrasts on the entries in $\boldsymbol{\gamma}$ are of interest in both models, those in B_γ. We will proceed to show that, with a mild restriction, the optimal designs for the estimation of a set of attribute effects when the Bradley–Terry model is used are also optional for the estimation of the same set of attribute effects independently of the contrasts on the entries in $\boldsymbol{\psi}$ when the Davidson–Beaver position effects model is used.

If we compare the generic diagonal entry of the $\Lambda(\boldsymbol{\pi}_0)$ matrix when the Bradley–Terry model is used, denoted by $\Lambda(\boldsymbol{\pi}_0)_{\text{B–T}}$, to the generic diagonal entries of $\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})$ when the Davidson–Beaver model is used, denoted by $\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\text{D–B}}$, we notice that

$$\frac{(\psi_1+\psi_2)^2}{\psi_1\psi_2}(\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\text{D–B}})_{ii} = 4(\Lambda(\boldsymbol{\pi}_0)_{\text{B–T}})_{ii} = \frac{1}{N}\sum_{i_2\neq i}\left(n_{(i,i_2)} + n_{(i_2,i)}\right), \tag{4.7}$$

where order is considered important. Similarly, if we compare the off–diagonal entries of the Λ matrices for the two models, we find that

$$\frac{(\psi_1+\psi_2)^2}{\psi_1\psi_2}(\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\text{D–B}})_{ij} = 4(\Lambda(\boldsymbol{\pi}_0)_{\text{B–T}})_{ij} = -\frac{n_{(i,j)} + n_{(j,i)}}{N}, \tag{4.8}$$

where order is considered important.

We also need to look at conditions that allow the attribute effects and the position main effect to be estimated independently. This will require the $(1,2)$ block, and therefore the $(2,1)$ block, of the information matrix for, $C_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi})$, to be 0. Before we can find the conditions for which these blocks are equal to 0, we need to find the expression for the entries in these blocks.

Recall that the $(1,2)$ block of $C(\boldsymbol{\pi}_0, \boldsymbol{\psi})$ is equal to

$$C_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = B_\gamma \Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) B_\psi^T.$$

A generic entry in the matrix obtained when multiplying the first two of these matrices together is

$$\begin{aligned}(B_\gamma \Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}))_{ja} &= \sum_{i=1}^{T} B_{ji} \times \frac{\psi_b}{(\psi_1+\psi_2)^2}(\lambda_{T_i \text{ in position } a} - \lambda_{T_i \text{ in position } b}) \\ &= \frac{\psi_b}{(\psi_1+\psi_2)^2}\sum_{i=1}^{T} B_{ji}(\lambda_{T_i \text{ in position } a} - \lambda_{T_i \text{ in position } b}),\end{aligned}$$

where $b \neq a$.

While it is obvious that $C_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = 0$ if all items that appear in the experiment appear equally often in both positions, this condition proves to be too restrictive. This constraint allows us to estimate all higher order effects independently of the contrasts of the ψ_a. Usually we are only interested in the estimation of the main effects of the attributes, and perhaps also the two–factor interactions between attributes, independently of the position main effect.

We now prove two lemmas which give conditions for $C_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = 0$, one when the attribute main effects and the position main effect are of interest, and the other when the main effects plus two–factor interactions of the attributes and the position main effect are of interest.

■ **LEMMA 4.4.1.**

The information matrix for the estimation of the main effects of the attributes and the position main effect is block diagonal if each of the levels for each attribute appears in each position equally often.

Proof. For an attribute main effect, every item with the same level for the corresponding attribute will have the same contrast coefficient. Then, for the j^{th} contrast corresponding to a component of the main effect of attribute q,

$$(B_\gamma \Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}))_{ja} = \frac{\psi_b}{(\psi_1+\psi_2)^2}\sum_{x=1}^{\ell_q} B_{jx}(\lambda_{\text{att } q=x \text{ in pos } a} - \lambda_{\text{att } q=x \text{ in pos } b}),$$

where $b \neq a$. Then $(B_\gamma \Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}))_{ja} = 0$ if $\lambda_{\text{att } q=x \text{ in pos } a} - \lambda_{\text{att } q=x \text{ in pos } b} = 0$ for all attribute levels $0 \leq x \leq \ell_q - 1$ and $b \neq a$. It follows that $B_\gamma \Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = \mathbf{0}$ if, for each attribute, each attribute level appears in both positions equally often. If this is the case, then the information matrix for the estimation of main effects and the position main effect is block diagonal. □

■ **LEMMA 4.4.2.**

The information matrix for the estimation of the main effects plus two–factor interactions of the attributes and the position main effect is block diagonal if, for each pair of attributes, each pair of attribute levels appears equally often in both positions of the choice set.

Proof. This proof follows the same lines as the proof of Lemma 4.4.1. The contrast coefficients corresponding to items with the same pair of levels will be the same. Then for the j^{th} contrast, corresponding to a component of the two–factor interaction between attributes q_1 and q_2,

$$(B_\gamma \Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}))_{ja} = \frac{\psi_b}{(\psi_1+\psi_2)^2} \sum_{x_1=1}^{\ell_{q_1}} \sum_{x_2=1}^{\ell_{q_2}} B_{j(x_1 x_2)} (\lambda_{q_1=x_1, q_2=x_2 \text{ in pos } a} - \lambda_{q_1=x_1, q_2=x_2 \text{ in pos } b}),$$

where $\lambda_{q_1=x_1,q_2=x_2 \text{ in pos } a} = \lambda_{\text{att } q_1=x_1,\text{ att } q_2=x_2 \text{ in pos } a}$. Then $(B_\gamma \Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}))_{ja} = 0$ if, for all x_1 and x_2,

$$\lambda_{q_1=x_1,q_2=x_2 \text{ in pos } a} - \lambda_{q_1=x_1,q_2=x_2 \text{ in pos } b} = 0$$

for $b \neq a$. This will also result in $(B_\gamma \Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}))_{ja} = 0$, where contrast j corresponds to a main effect of either attribute q_1 or attribute q_2, as each level of the attribute must appear in both positions equally often if each pair of attribute levels appears in each position equally often. It follows that $B_\gamma \Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = \mathbf{0}$ if, for each pair of attributes, each pair of attribute levels appears equally often in both of the positions in the choice set. Then the information matrix for the estimation of the main effects plus two–factor interactions and the position main effect is block diagonal. □

Now that we have compared the information matrices for the estimation of the Bradley–Terry model and the estimation of the Davidson–Beaver position effects model, as well as the conditions that make the estimation of position effects independent of the estimation of contrasts of the entries in $\boldsymbol{\gamma}$, we can compare optimality results for designs for these two models. We will use the D–optimality criterion as defined in Section 1.3.1.

■ **THEOREM 4.4.3.**

Consider a particular set of contrasts of the elements in $\boldsymbol{\gamma}$ and some constant, but unknown, values for the elements in $\boldsymbol{\psi}$. Let ξ_{OPT} be the D–optimal design for the estimation of a set of contrasts of the entries in $\boldsymbol{\gamma}$ using the Bradley–Terry model over the set of competing designs $\mathfrak{X}$. Then ξ_{OPT} is also D–optimal over $\mathfrak{X}$ for the estimation of the same set of contrasts of the entries on $\boldsymbol{\gamma}$ and a contrast of the elements in $\boldsymbol{\psi}$ that is constant across the designs in $\mathfrak{X}$, provided that the $(1,2)$ and $(2,1)$ blocks of the partitioned information matrix are $\mathbf{0}$. □

Proof. We begin by letting B be the block diagonal matrix

$$B = \begin{bmatrix} B_\gamma & \mathbf{0} \\ \mathbf{0} & B_\psi \end{bmatrix},$$

containing the contrasts of the entries in $\boldsymbol{\gamma}$ that are of interest, and the position main effect. All of these contrasts will be constant over the class of competing designs. Then the information matrix for the estimation of the contrasts in B_γ and B_ψ, assuming that the conditions of either Lemma 4.4.1 or Lemma 4.4.2 are satisfied, is

$$\begin{aligned} C(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\text{D–B}} &= B\Lambda(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\text{D–B}}B^T \\ &= \begin{bmatrix} B_\gamma\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\text{D–B}}B_\gamma^T & \mathbf{0} \\ \mathbf{0} & B_\psi\Lambda_{\psi\psi}(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\text{D–B}}B_\psi^T \end{bmatrix}. \end{aligned}$$

Equations 4.7 and 4.8 show that

$$\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\text{D–B}} = \frac{4\psi_1\psi_2}{(\psi_2+\psi_2)^2}\Lambda(\boldsymbol{\pi}_0)_{\text{B–T}},$$

so by substitution we obtain

$$C(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\text{D–B}} = \begin{bmatrix} \frac{4\psi_1\psi_2}{(\psi_2+\psi_2)^2}B_\gamma\Lambda(\boldsymbol{\pi}_0)_{\text{B–T}}B_\gamma^T & \mathbf{0} \\ \mathbf{0} & B_\psi\Lambda_{\psi\psi}(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\text{D–B}}B_\psi^T \end{bmatrix}.$$

The information matrix for the estimation of the set of contrasts in B_γ when the Bradley–Terry model is used is

$$C(\boldsymbol{\pi}_0)_{\text{B–T}} = B_\gamma\Lambda(\boldsymbol{\pi}_0)_{\text{B–T}}B_\gamma^T.$$

Then we may express $C(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\text{D–B}}$ in terms of $C(\boldsymbol{\pi}_0)_{\text{B–T}}$ and $\boldsymbol{\psi}$, which gives

$$C(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\text{D–B}} = \begin{bmatrix} \frac{4\psi_1\psi_2}{(\psi_2+\psi_2)^2}C(\boldsymbol{\pi}_0)_{\text{B–T}} & \mathbf{0} \\ \mathbf{0} & C_\psi(\boldsymbol{\psi}) \end{bmatrix}, \tag{4.9}$$

where

$$C_\psi(\boldsymbol{\psi}) = B_\psi\Lambda_{\psi\psi}(\boldsymbol{\pi}_0,\boldsymbol{\psi})B_\psi^T,$$

and is a function of the entries in $\boldsymbol{\psi}$ only.

Since ξ_{OPT} is the D–optimal design for the estimation of the set of contrasts in B_γ when the Bradley–Terry model is used, we have

$$\det((C(\boldsymbol{\pi}_0)_{\text{B–T}})_\xi) \leq \det((C(\boldsymbol{\pi}_0)_{\text{B–T}})_{\xi_{\text{OPT}}})$$

for all $\xi \in \mathfrak{X}$, by the definition of D–optimality, as given in Section 1.3.1. Using Equation 4.9 we see that

$$\begin{aligned} \det(C(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\text{D–B}}) &= \det\left(\frac{4\psi_1\psi_2}{(\psi_2+\psi_2)^2}C(\boldsymbol{\pi}_0)_{\text{B–T}}\right)\times\det(C_\psi(\boldsymbol{\psi})) \\ &= \frac{(4\psi_1\psi_2)^p}{(\psi_2+\psi_2)^{2p}}\times\det(C(\boldsymbol{\pi}_0)_{\text{B–T}})\times\det(C_\psi(\boldsymbol{\psi})). \end{aligned}$$

Since p and $\boldsymbol{\psi}$ are constant across the set of competing designs, so is $\det C_\psi(\boldsymbol{\psi})$, and thus the efficiency of an arbitrary design ξ when compared to the design ξ_{OPT} when using the Davidson–Beaver position effects model is

$$\begin{aligned} D_{\text{eff}}(\xi,\xi_{\text{OPT}}) &= \left(\frac{\det((C(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\text{D–B}})_\xi)}{\det((C(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\text{D–B}})_{\xi_{\text{OPT}}})}\right)^{1/(p+1)} \\ &= \left(\frac{\frac{(4\psi_1\psi_2)^p}{(\psi_2+\psi_2)^{2p}}\det((C(\boldsymbol{\pi}_0)_{\text{B–T}})_\xi)\times\det(C_\psi(\boldsymbol{\psi}))}{\frac{(4\psi_1\psi_2)^p}{(\psi_2+\psi_2)^{2p}}\det((C(\boldsymbol{\pi}_0)_{\text{B–T}})_{\xi_{\text{OPT}}})\times\det(C_\psi(\boldsymbol{\psi}))}\right)^{1/(p+1)} \\ &= \left(\frac{\det((C(\boldsymbol{\pi}_0)_{\text{B–T}})_\xi)}{\det((C(\boldsymbol{\pi}_0)_{\text{B–T}})_{\xi_{\text{OPT}}})}\right)^{1/(p+1)} \\ &\leq 1. \end{aligned}$$

Therefore, by the definition of D–optimality, ξ_{OPT} is also a D–optimal design for the estimation of the contrasts in B_γ and B_ψ when the Davidson–Beaver position effects model is used, assuming the conditions in the statement of Lemmas 4.4.1 and 4.4.2 are satisfied. □

We now consider an example of the relationship between these two models, and compare some designs.

■ **EXAMPLE 4.4.1.**

Recall the experiment introduced in Example 2.0.12 and design introduced in Example 4.1.2. In Example 4.3.1 we found the information matrix for the estimation of main effects and the position main effect when the Davidson–Beaver position effects model is used. Now we will find the information matrix for the estimation of the main effects only using the same design when the Bradley–Terry model is used. The contrast matrix for the estimation of main effects is

$$B = \frac{1}{2}\begin{bmatrix} -1 & -1 & 1 & 1 \\ -1 & 1 & -1 & 1 \end{bmatrix}.$$

From El-Helbawy and Bradley [1978] we know that the information matrix for the estimation of $\boldsymbol{\gamma}$ when the Bradley–Terry model is used, under the assumption of the null hypothesis of equal merits, is

$$\Lambda(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\text{B–T}} = \frac{1}{8}\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}.$$

From Example 4.2.2, it is clear that

$$\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\text{D–B}} = \frac{4\psi_1\psi_2}{(\psi_2+\psi_2)^2}\Lambda(\boldsymbol{\pi}_0)_{\text{B–T}}.$$

It follows that the information matrix for the estimation of main effects only when the Bradley–Terry model is used is

$$C(\boldsymbol{\pi}_0)_{\text{B–T}} = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{4} \end{bmatrix},$$

and we see that

$$C(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\text{D–B}} = \begin{bmatrix} \frac{4\psi_1\psi_2}{(\psi_2+\psi_2)^2}C(\boldsymbol{\pi}_0)_{\text{B–T}} & 0 \\ 0 & \frac{1}{2\psi_1\psi_2} \end{bmatrix}.$$

Taking determinants of both $C(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\text{D–B}}$ and $C(\boldsymbol{\pi}_0)_{\text{B–T}}$ gives

$$\begin{aligned} \det(C(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\text{D–B}}) &= \frac{\psi_1\psi_2}{2\,(\psi_1+\psi_2)^{\,4}}, \\ \det(C(\boldsymbol{\pi}_0)_{\text{B–T}}) &= \frac{1}{16}. \end{aligned}$$

We are estimating $p = 2$ contrasts on the entries in $\boldsymbol{\gamma}$, which gives

$$\begin{aligned} \frac{(4\psi_1\psi_2)^p}{(\psi_1+\psi_2)^{2p}} \times \det(C(\boldsymbol{\pi}_0)_{\text{B–T}}) \times \det(C_\psi(\boldsymbol{\psi})) &= \frac{(4\psi_1\psi_2)^2}{(\psi_1+\psi_2)^4} \times \frac{1}{16} \times \frac{1}{2\gamma_1\gamma_2} \\ &= \frac{\psi_1\psi_2}{2(\psi_1+\psi_2)^4} \\ &= \det(C(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\text{DAV}}), \end{aligned}$$

which is consistent with the findings in Theorem 4.4.3. □

We can use this theorem to apply the results that we know about optimal designs for the Bradley–Terry model to the Davidson–Beaver position effects model. First, we apply Theorem 4.4.3 for a 2^k factorial experiment presented in Theorem 1.3.1.

■ **COROLLARY 4.4.4.**
Let ξ be the design that contains all distinct pairs that differ in the levels of each attribute in a 2^k paired comparisons experiment. Then when the rows of B_γ correspond to the k main effects, and each attribute level appears in both positions equally often, the design will be D–optimal for the estimation of the main effects of the attributes and the position main effect when the Davidson–Beaver position effects model is used. □

Proof. By Theorem 1.3.1, the design in the statement of the theorem is D–optimal for the estimation of the attribute main effects when the Bradley–Terry model is used. By Lemma 4.4.1, the design will have a block diagonal information matrix. It follows from Theorem 4.4.3 that this design is D–optimal for the estimation of the attribute main effects and the position main effect when the Davidson–Beaver position effects model is used. □

We can use this corollary to find an optimal design for the estimation of the main effects and the position main effect for the experiment in our example.

■ **EXAMPLE 4.4.2.**
Consider the 2^2 experiment introduced in Example 2.0.12. In this experiment we have $t = 4$ possible items. There are four ordered pairs of items, shown in Table 4.2, that differ in all $k = 2$ attributes, and give each attribute level appearing in each position twice. Then the design with these four ordered pairs is optimal for the estimation of main effects and the position main effect when the Davidson–Beaver position effects model is used. □

Option 1	**Option 2**
0 0	1 1
0 1	1 0
1 0	0 1
1 1	0 0

Table 4.2: Optimal design for main effects and the position main effect.

We can also use Theorem 4.4.3 to extend the result on the optimal design for a 2^k factorial for the estimation of the main effects plus two–factor interactions and the position main effect given in Theorem 1.3.2 .

■ **COROLLARY 4.4.5.**
The D–optimal design for the estimation of the main effects plus two–factor interactions of the attributes and the position main effect for a 2^k paired comparisons experiment, when all other effects are assumed zero and the Davidson–Beaver potion effects model is used, is given by

$$a_{k,i} = \begin{cases} 2^{k-1}\binom{k}{(k+1)/2}^{-1}, & \text{if } i = \frac{k+1}{2}, \\ 0, & \text{otherwise.} \end{cases}$$

if k is odd and

$$a_{k,i} = \begin{cases} 2^{k-1}\binom{k}{k/2}^{-1}, & \text{if } i = \frac{k}{2} \text{ or } \frac{k}{2}+1, \\ 0, & \text{otherwise.} \end{cases}$$

if k is even, provided that, for each pair of attributes, each pair of attribute levels appear equally often in both positions. □

Proof. By Theorem 1.3.2, the design in the statement of the theorem is D–optimal for the estimation of the main effects plus two–factor interactions of the attributes when the Bradley–Terry model is used. By Lemma 4.4.2, this design will also have a block diagonal information matrix. It follows from Theorem 4.4.3 that this design is D–optimal for the estimation of the main effects plus two–factor interactions of the attributes and the position main effect when the Davidson–Beaver position effects model is used. □

We can use this corollary to find an optimal design for the estimation of the main effects plus two–factor interactions and the position main effect for the experiment in our examples.

■ **EXAMPLE 4.4.3.**

Consider again the 2^2 experiment introduced in Example 2.0.12. We are now interested in the estimation of the main effects plus two–factor interactions of the attributes and the position main effect. Since $k = 2$ is even the D–optimal design for the estimation of these effects is given by

$$a_{2,i} = \begin{cases} \frac{1}{2} \times \binom{2}{1}^{-1}, & \text{if } i = 1 \text{ or } 2, \\ 0, & \text{otherwise.} \end{cases}$$

This is the design with all ordered pairs of distinct items, as shown in Table 4.3. □

Option 1	**Option 2**	**Option 1**	**Option 2**
0 0	0 1	1 0	0 0
0 1	0 0	1 1	0 1
1 0	1 1	0 0	1 1
1 1	1 0	0 1	1 0
0 0	1 0	1 0	0 1
0 1	1 1	1 1	0 0

Table 4.3: Optimal design for main effects, two–factor interactions, and the position main effect.

Finally, we can use Theorem 4.4.3 to extend the results relating to the optimal design for the general factorial as given in Theorem 1.3.3.

■ **COROLLARY 4.4.6.**

Consider an $\ell_1 \times \ldots \times \ell_q$ factorial paired comparisons experiment. Assuming that there are no interactions present, and B_h contains the contrast coefficients for the attribute main effects, then the design consisting of all pairs where the options differ in all of the attributes, and where each attribute level appears equally often in both positions, will be D–optimal for the estimation of the main effects of the attributes and the position main effect when the Davidson–Beaver position effects model is used. □

Proof. By Theorem 1.3.3, the design in the statement of the theorem is D–optimal for the estimation of the attribute main effects in B_h when the Bradley–Terry model is used. By Lemma 4.4.1, the design will also have a block diagonal information matrix. It follows from Theorem 4.4.3 that this design is D–optimal for the estimation of the main effects of the attributes and the position main effect when the Davidson–Beaver position effects model is used. □

We now consider an example of how this result can be used to find optimal designs for the estimation of main effects when using the Davidson–Beaver model.

■ **EXAMPLE 4.4.4.**
Let us consider the 3^2 experiment introduced in Example 2.4.4. An optimal design for the estimation of the main effects of the attributes and the position main effect is given in Table 4.4. Notice that each pair of items that differs in each of the attributes appears, and that each level in each of the attributes appears 6 times in each position of the choice set, satisfying the criteria for Corollary 4.4.6. □

Option 1	Option 2	Option 1	Option 2
0 0	1 1	0 0	2 2
0 1	1 2	0 1	2 0
0 2	1 0	0 2	2 1
1 0	2 1	1 0	0 2
1 1	2 2	1 1	0 0
1 2	2 0	1 2	0 1
2 0	0 1	2 0	1 2
2 1	0 2	2 1	1 0
2 2	0 0	2 2	1 1

Table 4.4: Optimal design for main effects and the position main effect, $\ell_1 = \ell_2 = 3$.

4.5 Simulations of the Davidson–Beaver Model

In this section we consider the performance of the Davidson–Beaver position effects model under various model assumptions by carrying out a number of simulation studies. We assume that $k = 2$, $\ell_1 = \ell_2 = 2$ and $m = 2$ throughout. We consider two sets of values for the parameters. In the first we assume that both main effects parameters, τ_1 and τ_2, are equal to 0 and the position main effect parameter $\psi_{\rm L} = -0.2$, and in the second set we assume that $\tau_1 = 1$ and $\tau_2 = -1$ but $\psi_{\rm L} = -0.2$ still.

We find the locally optimal design for each set of values and compare the performance of each design with both sets of parameter values. The design in Table 4.2 is optimal for the estimation of the main effects of the attributes plus the position main effect when $\tau_1 = \tau_2 = 0$ and $\psi_{\rm L} = -0.2$, as shown in Example 4.4.2. By an exhaustive search of the $2^{12} - 1 = 4095$ possible designs, the design in Table 4.5 is optimal for the estimation of the main effects of the attributes plus the position main effect when $\tau_1 = 1$, $\tau_2 = -1$, and $\psi_{\rm L} = -0.2$. This exhaustive search is illustrated in Figure 4.1, where the x–coordinate corresponds to the design index, and the y–coordinate is the determinant of the information matrix for that design when $\tau_1 = 1$, $\tau_2 = -1$, and $\psi_L = -0.2$. The determinants of the information matrix for the designs in Tables 4.2 and 4.5 are labelled in Figure 4.1.

We first assume that $\tau_1 = \tau_2 = 0$, and $\psi_L = -0.2$ and compare the simulated distributions of the parameter estimates when the designs in Tables 4.2 and 4.5 are used in turn. Each simulation is modelled using the simulated responses from 150 respondents, and each boxplot displays the distribution of the estimates from 1000 such simulations. Figures 4.2(a) and (b) show the distributions of the parameter estimates when the designs in Tables 4.2 and 4.5, respectively, are used. Summary statistics for both simulations are provided in Table 4.6. We see that, in each of the simulations, the simulated parameter estimates are unbiased and symmetrically distributed. We see that, in this case, the variances of the parameter estimates for the design in Table 2.6 is larger, illustrating the optimality of the design in Table 4.2.

We now consider the performance of these two designs when $\tau_1 = 1, \tau_2 = -1$, and $\psi_L = -0.2$. Figures 4.3(a) and (b) show the distributions of the parameter estimates when the designs in Tables 4.2 and 4.5, respectively, are used. Summary statistics for both simulations are provided in Table 4.7. We see that, for both designs, the distribution of the parameter estimates seem to be unbiased. In this case, we see that the variance of the parameter estimates for the design in Table 4.5 is now smaller, illustrating the optimality of this design. The selection probabilities for the design in Table 4.2 when $\tau_1 = 1$, $\tau_2 = -1$, and $\psi_L = -0.2$ are given in Table 4.8.

Next, we simulate the effect of changing the coefficient of the position main effect on the distributions of the parameter estimates when we let $\tau_1 = 1$ and $\tau_2 = -1$, and the design in Table 4.5 is used. Figures 4.4(a) and (b) give the simulated distributions of the parameter estimates when the coefficient of the position main effect is 0 and -0.4, respectively. Summary statistics for both simulations are provided in Table 4.9. We see that in each of the simulations that the parameter estimates are unbiased, fairly symmetrically distributed and have similar variances.

We now compare the ability of four different designs to estimate the main effects plus the two–factor interaction of the attributes and the position main effect. The first two designs are those in Tables 4.2 and 4.5. The third design is the set of all ordered pairs of items, which is optimal for the estimation of the main effects plus the two–factor interaction of the attributes and the position main effect when $\tau_1 = \tau_2 = \tau_{12} = 0$, and $\psi_L = -0.2$, as shown in Example 4.4.3. This design is shown in Table 4.3. The final design, shown in Table 3.12, is locally optimal for

Option 1	**Option 2**
0 0	0 1
0 0	1 0
0 0	1 1
1 1	0 0

Table 4.5: Optimal design for main effects and the position main effect when $\tau_1 = 1$, $\tau_2 = -1$, and $\psi_L = -0.2$.

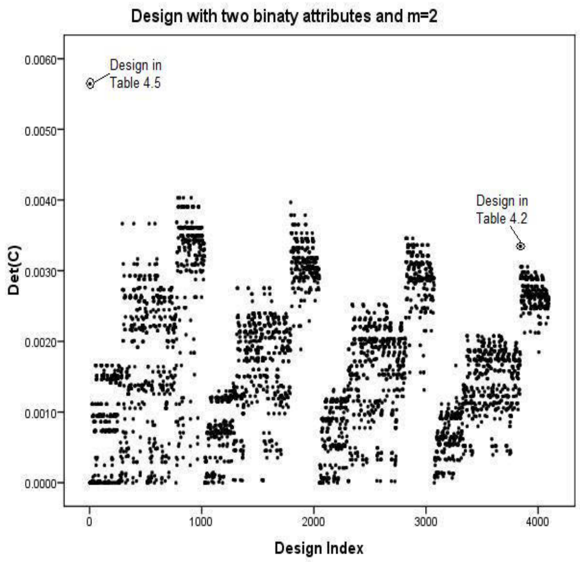

Figure 4.1: Exhaustive search for optimal design $\tau_1 = 1, \tau_2 = -1$, and $\psi_{\mathrm{L}} = -0.2$.

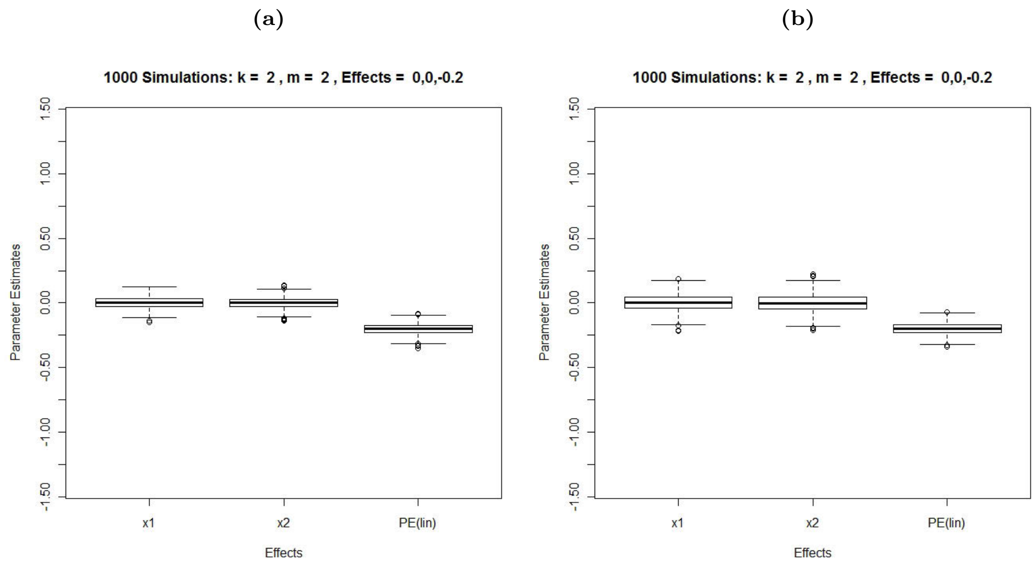

Figure 4.2: Simulation: $\tau_1 = 0$, $\tau_2 = 0$, and $\psi_{\mathrm{L}} = -0.2$.

Parameter	Simulated Mean (Standard Error)	Theoretical Variance	Simulated Variance	Simulated Skewness (Standard Error)
Design in Table 4.2				
τ_1	0.00266(0.00138)	0.00168	0.00192	−0.05424(0.07734)
τ_2	0.00004(0.00133)	0.00168	0.00176	0.14714(0.07734)
ψ_{L}	−0.20150(0.00132)	0.00167	0.00173	−0.10012(0.07734)
Design in Table 4.5				
τ_1	0.00247(0.00202)	0.00411	0.00409	−0.15819(0.07734)
τ_2	−0.00167(0.00207)	0.00411	0.00428	−0.02982(0.07734)
ψ_{L}	−0.20152(0.00143)	0.00185	0.00205	−0.00819(0.07734)

Table 4.6: Summary statistics for $\tau_1 = 0$, $\tau_2 = 0$, and $\psi_{\mathrm{L}} = -0.2$.

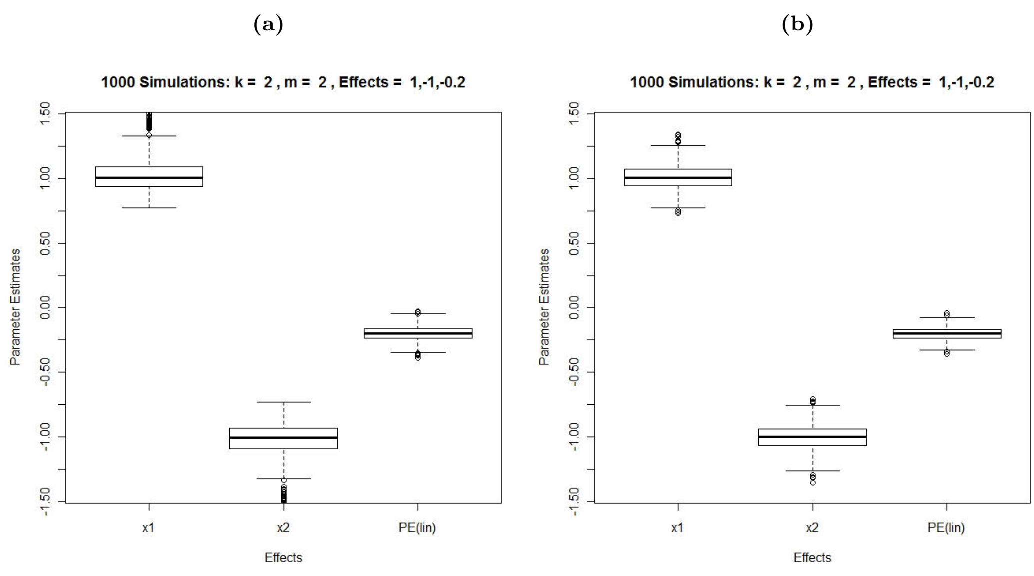

Figure 4.3: Simulation: $\tau_1 = 1$, $\tau_2 = -1$, and $\psi_{\mathrm{L}} = -0.2$.

Parameter	Simulated Mean (Standard Error)	Theoretical Variance	Simulated Variance	Simulated Skewness (Standard Error)
Design in Table 4.2				
τ_1	1.02515(0.00417)	0.00191	0.01735	0.86034(0.07734)
τ_2	−1.025238(0.00417)	0.00191	0.01739	−1.03217(0.07734)
ψ_L	−0.19870(0.00179)	0.00187	0.00320	−0.10202(0.07734)
Design in Table 4.5				
τ_1	1.00814(0.00294)	0.00429	0.00863	0.14050(0.07734)
τ_2	−1.00626(0.00307)	0.00437	0.00943	−0.15397(0.07734)
ψ_L	−0.20048(0.00157)	0.00189	0.00245	−0.09339(0.07734)

Table 4.7: Summary statistics for $\tau_1 = 1$, $\tau_2 = -1$, and $\psi_L = -0.2$.

Choice Set	$\mathbf{P(T_1\|(T_1, T_2))}$	$\mathbf{P(T_2\|(T_1, T_2))}$
(00, 11)	0.599	0.401
(01, 10)	0.168	0.832
(10, 01)	0.919	0.083
(11, 00)	0.599	0.401

Table 4.8: Selection probabilities for the design in Table 4.2, where $\tau_1 = 1$, $\tau_2 = -1$, and $\psi_L = -0.2$.

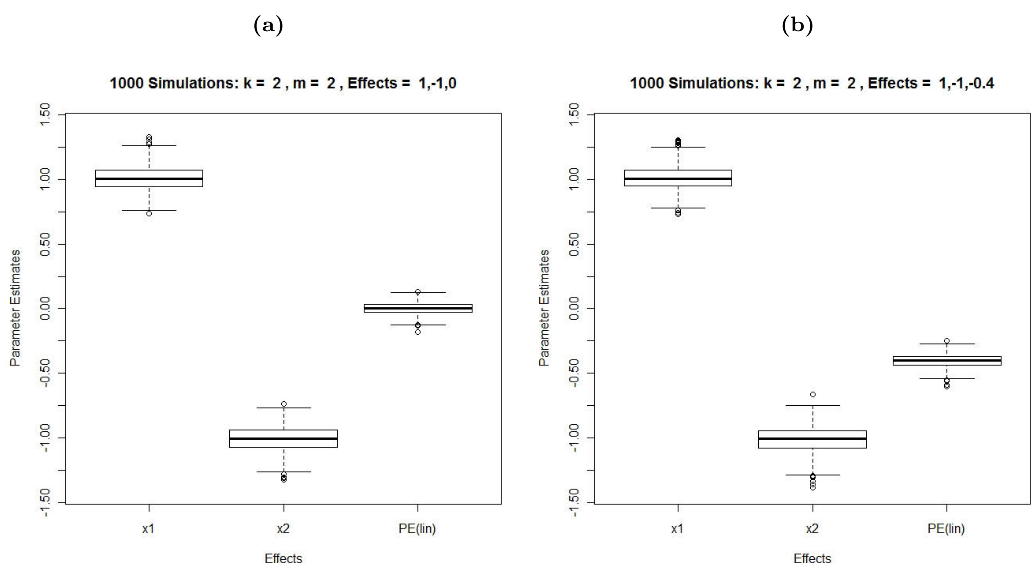

Figure 4.4: Simulation: $\tau_1 = 0.5$, $\tau_2 = -0.5$, and $\psi_L = 0$ (a), and -0.4 (b).

	Parameter	Simulated Mean (Standard Error)	Theoretical Variance	Simulated Variance	Simulated Skewness (Standard Error)
(a)	τ_1	1.00998(0.00299)	0.00429	0.00895	0.13724(0.07734)
	τ_2	−1.01125(0.00324)	0.00429	0.01050	−0.15104(0.07734)
	ψ_{L}	−0.40587(0.00174)	0.00189	0.00301	−0.19620(0.07734)
(b)	τ_1	1.00953(0.00300)	0.00437	0.00903	0.31702(0.07734)
	τ_2	−1.01020(0.00300)	0.00453	0.00902	−0.31558(0.07734)
	ψ_{L}	0.00129(0.00158)	0.00189	0.00248	−0.00143(0.07734)

Table 4.9: Summary statistics for $\tau_1 = 1$, $\tau_2 = -1$, and (a) $\psi_{\mathrm{L}} = 0$ and (b) $\psi_{\mathrm{L}} = -0.4$.

the estimation of the main effects plus the two–factor interaction of the attributes and position the main effect when $\tau_1 = 1, \tau_2 = -1, \tau_{12} = -0.25$, and $\psi_{\mathrm{L}} = -0.2$, found by using an exhaustive search.

We first consider the case where the two–factor interaction is negligible. We let the coefficients of the main effects be $\tau_1 = 1$ and $\tau_2 = -1$, and the coefficient of the position main effect be -0.2. Then Figures 4.5(a), (b), (c), and (d) give the distributions of the parameter estimates when the designs in Table 4.2, Table 4.5, Table 4.3, and Table 4.10 are used. Summary statistics for all four of the simulations are provided in Table 4.11.

We see that the designs in Tables 4.3 and 4.10 give unbiased and close to symmetric distributions with similar variances. The design in Table 4.5 gives slightly biased but reasonably symmetrically distributed parameter estimates with a larger variance than the designs in Table 4.3 and 4.10. The design in Table 4.2 can not estimate the two–factor interaction at all.

Finally, we consider the estimation of a non–zero interaction effect. If we let the coefficients of the main effects be $\tau_1 = 1$ and $\tau_2 = -1$, the coefficient of the position main effect be -0.2 as before, and we let the coefficient of the interaction effect be $\tau_{12} = -0.25$, then the selection probabilities for the design in Table 4.3 are given in Table 4.12. Figures 4.6(a), (b), (c), and (d) give the simulated distributions of the parameter estimates when the designs in Table 4.2, Table 4.5, Table 4.3, and Table 4.10 are used. Summary statistics for all four of the simulations are

Option 1	Option 2	Option 1	Option 2
0 1	0 0	1 1	0 1
1 0	1 1	0 0	1 1
1 1	1 0	0 1	1 0
0 0	1 0	1 0	0 1
0 1	1 1	1 1	0 0
1 0	0 0		

Table 4.10: Optimal design for main effects plus two–factor interactions of the attributes and the position main effect when $\tau_1 = 1$, $\tau_2 = -1$, $\tau_{12} = -0.25$, and $\psi_{\mathrm{L}} = -0.2$.

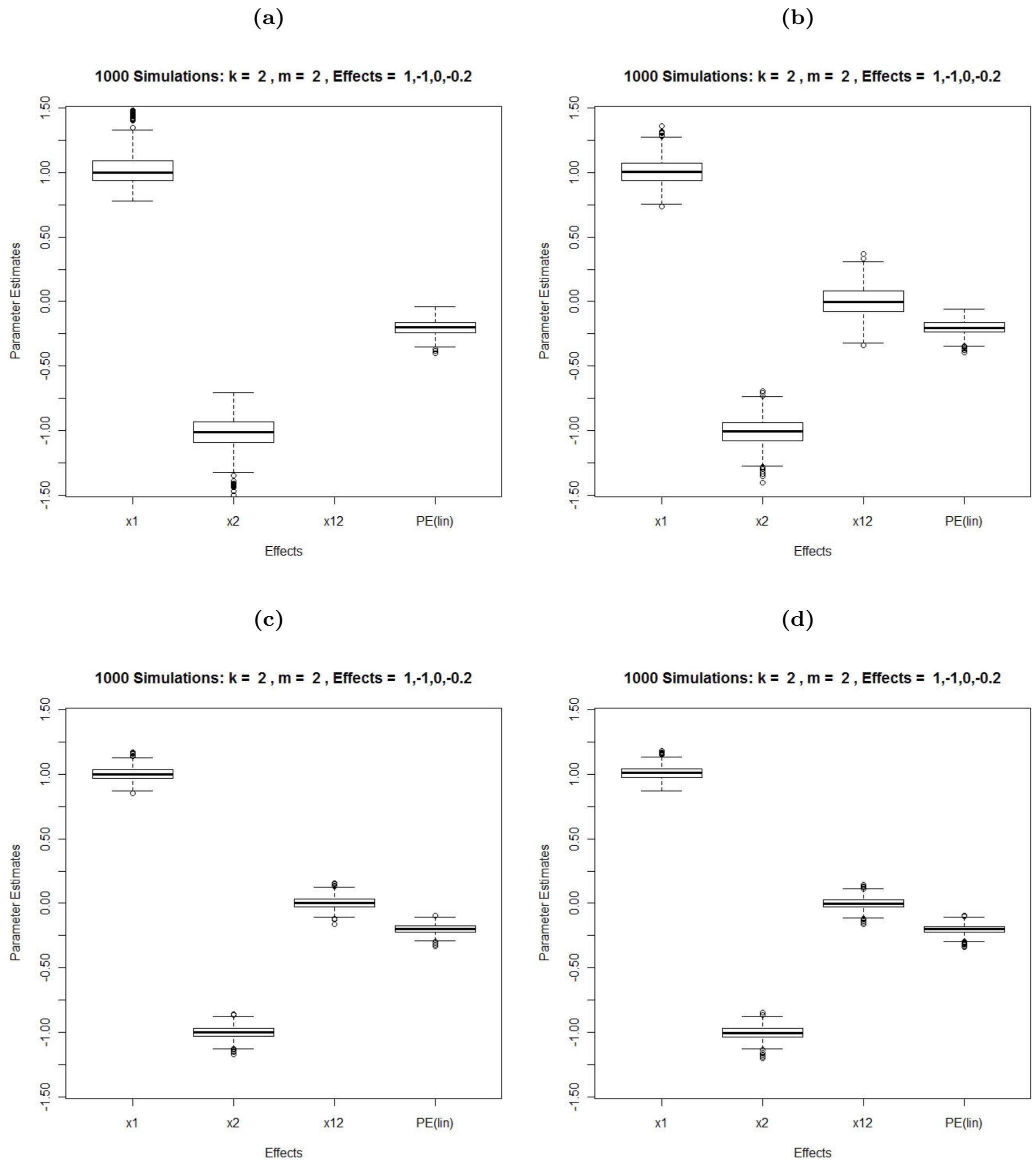

Figure 4.5: Simulation: Estimating main effects and the position main effect, designs in (a) Table 4.2, (b) Table 4.5, (c) Table 4.3, and (d) Table 4.10.

Parameter	Simulated Mean (Standard Error)	Theoretical Variance	Simulated Variance	Simulated Skewness (Standard Error)
Design in Table 4.2				
τ_1	1.02367(0.00385)	0.00191	0.01481	0.82857(0.07734)
τ_2	−1.02281(0.00389)	0.00191	0.01509	−0.73502(0.07734)
τ_{12}	Not Estimable			
ψ_{L}	−0.20072(0.00186)	0.00187	0.00343	−0.08013(0.07734)
Design in Table 4.5				
τ_1	1.00942(0.00310)	0.00428	0.00963	0.19616(0.07734)
τ_2	−1.01090(0.00325)	0.00428	0.01059	−0.20917(0.07734)
τ_{12}	0.00054(0.00364)	0.00755	0.01328	−0.00248(0.07734)
ψ_{L}	−0.20118(0.00180)	0.00335	0.00325	−0.30448(0.07734)
Design in Table 4.3				
τ_1	1.00121(0.00153)	0.00087	0.00233	0.29920(0.07734)
τ_2	−1.00138(0.00157)	0.00086	0.00245	−0.25802(0.07734)
τ_{12}	0.00323(0.00139)	0.00086	0.00193	0.03527(0.07734)
ψ_{L}	−0.19889(0.00113)	0.00057	0.00128	−0.21674(0.07734)
Design in Table 4.10				
τ_1	1.00914(0.00159)	0.00103	0.00254	0.18692(0.07734)
τ_2	−1.00492(0.00156)	0.00101	0.00243	−0.12623(0.07734)
τ_{12}	−0.00131(0.00137)	0.00086	0.00189	0.00943(0.07734)
ψ_{L}	−0.20199(0.00118)	0.00064	0.00140	−0.20892(0.07734)

Table 4.11: Summary statistics for $\tau_1 = 1, \tau_2 = -1, \tau_{12} = 0$, and $\psi_{\text{L}} = -0.2$.

provided in Table 4.13.

Once again, we see that the designs in Tables 4.3 and 4.10 give unbiased and close to symmetric distributions with similar variances. The design in Table 4.5 again gives slightly biased but reasonably symmetrically distributed parameter estimates with a larger variance than the designs in Table 4.3 and 4.10. The design in Table 4.2 still can not be used to estimate the two–factor interaction at all.

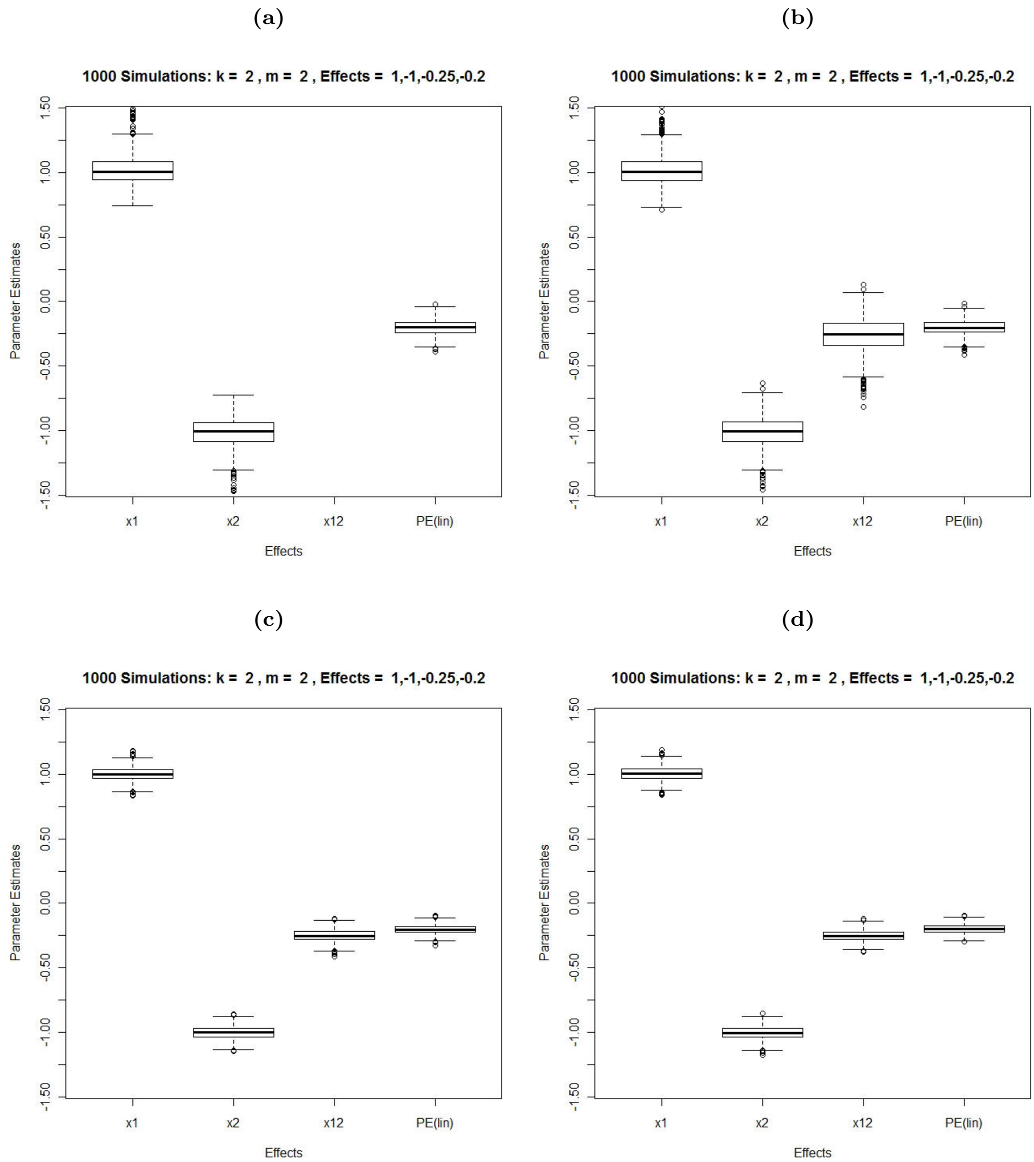

Figure 4.6: Simulation: Estimating main effects, two–factor interactions and the position main effect, designs in (a) Table 4.2, (b) Table 4.5, (c) Table 4.3, and (d) Table 4.10.

Choice Set	$\mathbf{P(T_1 \vert (T_1, T_2))}$	$\mathbf{P(T_2 \vert (T_1, T_2))}$
$(00, 01)$	0.948	0.052
$(01, 00)$	0.109	0.891
$(10, 11)$	0.870	0.130
$(11, 10)$	0.250	0.750
$(00, 10)$	0.250	0.750
$(01, 11)$	0.109	0.891
$(10, 00)$	0.870	0.130
$(11, 01)$	0.948	0.052
$(00, 11)$	0.599	0.401
$(01, 10)$	0.027	0.973
$(10, 01)$	0.988	0.012
$(11, 00)$	0.599	0.401

Table 4.12: Selection probabilities for the design in Table 4.2 when $\tau_1 = 1$, $\tau_2 = -1$, $\tau_{12} = -0.25$, and $\psi_{\mathrm{L}} = -0.2$.

Parameter	Simulated Mean (Standard Error)	Theoretical Variance	Simulated Variance	Simulated Skewness (Standard Error)
Design in Table 4.2				
τ_1	1.02376(0.00382)	0.00191	0.01457	0.92721(0.07734)
τ_2	−1.02326(0.00384)	0.00191	0.01471	−0.88053(0.07734)
τ_{12}	Not Estimable			
ψ_{L}	−0.20173(0.00185)	0.00187	0.00340	−0.13214(0.07734)
Design in Table 4.5				
τ_1	1.01627(0.00372)	0.00449	0.01382	0.65400(0.07734)
τ_2	−1.01545(0.00380)	0.00449	0.01441	−0.61478(0.07734)
τ_{12}	−0.26113(0.00412)	0.00779	0.01700	−0.44014(0.07734)
ψ_{L}	−0.20197(0.00184)	0.00330	0.00340	−0.06620(0.07734)
Design in Table 4.3				
τ_1	1.00329(0.00164)	0.00092	0.00269	0.02212(0.07734)
τ_2	−1.00288(0.00159)	0.00092	0.00254	−0.10721(0.07734)
τ_{12}	−0.25200(0.00144)	0.00090	0.00207	−0.13281(0.07734)
ψ_{L}	−0.20179(0.00112)	0.00060	0.00126	0.10934(0.07734)
Design in Table 4.10				
τ_1	1.00539(0.00165)	0.00107	0.00271	0.01579(0.07734)
τ_2	−1.00556(0.00163)	0.00107	0.00266	−0.16227(0.07734)
τ_{12}	−0.25139(0.00141)	0.00090	0.00200	0.03671(0.07734)
ψ_{L}	−0.20090(0.00112)	0.00067	0.00125	0.06408(0.07734)

Table 4.13: Summary statistics for $\tau_1 = 1, \tau_2 = -1, \tau_{12} = -0.25$, and $\psi_{\text{L}} = -0.2$.

4.A Proof that $\mathbf{j_L^T z} = 0$ for the Davidson–Beaver Position Effects Model

We begin by recalling that

$$z_i = \sum_{i_2 \neq i} w_{i|ii_2} - \frac{sn_{ii_2}\widehat{\psi_1}\widehat{\pi}_i}{\widehat{\psi_1}\widehat{\pi}_i + \widehat{\psi_2}\widehat{\pi_{i_2}}} + w_{i|i_2i} - \frac{sn_{i_2i}\widehat{\psi_2}\widehat{\pi}_i}{\widehat{\psi_1}\widehat{\pi_{i_2}} + \widehat{\psi_2}\widehat{\pi}_i}.$$

Now, the vector $\boldsymbol{z}$ contains the values for z_i for each possible item T_i. Then

$$\begin{aligned}
\boldsymbol{j}_L^T\boldsymbol{z} &= \sum_{i=1}^{t} z_i \\
&= \sum_{i_1 \neq i_2} w_{i_1|i_1i_2} + w_{i_1|i_2i_1} - \frac{sn_{i_1i_2}\widehat{\psi_1}\widehat{\pi_{i_1}}}{\widehat{\psi_1}\widehat{\pi_{i_1}} + \widehat{\psi_2}\widehat{\pi_{i_2}}} - \frac{sn_{i_2i_1}\widehat{\psi_2}\widehat{\pi_{i_1}}}{\widehat{\psi_1}\widehat{\pi_{i_2}} + \widehat{\psi_2}\widehat{\pi_{i_1}}} \\
&= \sum_C (w_{i_1|i_1i_2} + w_{i_2|i_1i_2}) - \sum_C \frac{sn_C\widehat{\psi_1}\widehat{\pi_{i_1}}}{\widehat{\psi_1}\widehat{\pi_{i_1}} + \widehat{\psi_2}\widehat{\pi_{i_2}}} + \frac{sn_C\widehat{\psi_2}\widehat{\pi_{i_2}}}{\widehat{\psi_1}\widehat{\pi_{i_1}} + \widehat{\psi_2}\widehat{\pi_{i_2}}} \\
&= \sum_C sn_C - \sum_C sn_C \\
&= 0,
\end{aligned}$$

as required.

Chapter 5

The Generalised Davidson–Beaver Position Effects Model

In Chapter 4, we found optimal designs for experiments using the Davidson–Beaver position effects model for choice sets of size 2. In Section 1.1, we introduced the MNL model as a generalisation of the Bradley–Terry model thus allowing for an arbitrary number of options in each choice set.

In this chapter we consider a generalisation of the MNL model to accommodate position effects. This generalisation is analogous to the generalisation of the Bradley–Terry model to obtain the Davidson–Beaver model. We will first set up the model, looking at probability density and likelihood functions and derive the information matrix for the estimation of a set of contrasts on the entries in $\boldsymbol{\gamma}$ and the position effects.

Once we have established some properties of this model, we will then use the information matrix to develop theory relating to the optimal design of experiments for this model. We conclude by looking at some relevant simulations.

5.1 Estimation of the generalised Davidson–Beaver position effects model

In this section we introduce a model that generalises the MNL model to incorporate position effects. We also derive the maximum likelihood estimates for this generalised model.

In order to estimate position effects for any choice set size, we need to introduce some additional parameters. The Davidson–Beaver position effects model incorporates position effects by multiplying the merit of item T_i, π_i, by a parameter ψ_a to reflect the effect of the item being presented in position a of the ordered choice set.

In general, we define $\psi_1, \psi_2, \ldots, \psi_m$ to be the parameters that measure the effect of an item being presented in positions $1, 2, \ldots, m$ of the choice set respectively. Then the probability of choosing an item T_i, which is presented in position a of the ordered choice set $C = (T_{i_1}, T_{i_2}, \ldots, T_{i_m})$, that is $T_i = T_{i_a}$, is

$$P(T_{i_a}|C) = \frac{\psi_a \pi_{i_a}}{\sum_{b=1}^{m} \psi_b \pi_{i_b}}.$$

To ensure identifiability, we assume that $\prod_{a=1}^{m} \psi_a = 1$. We now consider an example that applies these probabilities to the experiment considered in Example 3.1.1, where $m = 3$.

■ **EXAMPLE 5.1.1.**

Consider the experiment introduced in Example 3.1.1. If we assign merits π_{00}, π_{01}, and π_{10} to the items in the ordered choice set $C = (00, 01, 10)$, then

$$P(00|C) = \frac{\psi_1\pi_{00}}{\psi_1\pi_{00} + \psi_2\pi_{01} + \psi_3\pi_{10}},$$
$$P(01|C) = \frac{\psi_2\pi_{01}}{\psi_1\pi_{00} + \psi_2\pi_{01} + \psi_3\pi_{10}},$$

and

$$P(10|C) = \frac{\psi_3\pi_{10}}{\psi_1\pi_{00} + \psi_2\pi_{01} + \psi_3\pi_{10}}.$$

□

The next example looks at some of the values that $\psi_1, \ldots, \psi_m$ may take, and the types of position effect that they describe.

■ **EXAMPLE 5.1.2.**

Consider an experiment with $m = 5$ items in each choice set, and $\pi_1 = \ldots = \pi_L = 1$. Then $P(T_{i_a}|C) = \frac{\psi_a}{\sum_{b=1}^{m}\psi_b}$. Table 5.1 gives four different sets of values that $\boldsymbol{\psi}$ may take, and the probability of selection when all entries in $\boldsymbol{\pi}$ are equal to 1. The first two rows describe linear position effects, the first of which has the probability of selection increasing as the position moves from left to right, and the second has the probability of selection decreasing as the position moves from left to right. The final two rows describe a quadratic position effect. In the first of these, the probability of selection at the extremes of the choice set is higher that the probability of selection for the items in the middle positions. The second of these describes the opposite effect. □

$\boldsymbol{\psi}$	$\mathbf{P(T_{i_1}\|C)}$	$\mathbf{P(T_{i_2}\|C)}$	$\mathbf{P(T_{i_3}\|C)}$	$\mathbf{P(T_{i_4}\|C)}$	$\mathbf{P(T_{i_5}\|C)}$
$(0.64, 0.84, 1.04, 1.24, 1.44)$	0.123	0.162	0.200	0.238	0.277
$(1.44, 1.24, 1.04, 0.84, 0.64)$	0.277	0.238	0.200	0.162	0.123
$(1.31, 0.88, 0.74, 0.88, 1.31)$	0.256	0.172	0.145	0.172	0.256
$(0.74, 1.17, 1.32, 1.17, 0.74)$	0.144	0.228	0.257	0.228	0.144

Table 5.1: Selection probabilities for the experiment in Example 5.1.2.

Once again, we define indicator variables $\boldsymbol{w}$ to represent whether a particular decision was made by a particular respondent, α, or not. We let

$$w_{i|C\alpha} = \begin{cases} 1 & \text{if respondent } \alpha \text{ selects item } T_i \text{ from} \\ & \text{the ordered choice set } C = (T_{i_1}, T_{i_2}, \ldots, T_{i_m}), \\ 0 & \text{otherwise,} \end{cases}$$

where for a given choice set and respondent, only one of the w_is is equal to 1, depending on the respondent's choice. This is because there are no repeated ordered choice sets for any respondent, and no opt–out process. Then, for respondent α, the probability density function for the response to the ordered choice set $C = (T_{i_1}, T_{i_2}, \ldots, T_{i_m})$ is

$$f_{C\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}) = \frac{\prod_{a=1}^{m}(\psi_a\pi_{i_a})^{w_{i_a|C\alpha}}}{(\sum_{b=1}^{m}\psi_b\pi_{i_b})^{n_C}},$$

where n_C is an indicator variable which equals 1 if the ordered choice set C appears in the experiment, and 0 if it does not. For consistency we also let $w_{i|C\alpha} = 0$, for all items $T_i \in C$, if the ordered choice set C does not appear in the experiment. Given $f_{C\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi})$,

$$\ln(f_{C\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi})) = \sum_{a=1}^{m} w_{i_a|C\alpha}(\ln(\psi_a) + \ln(\pi_{i_a})) - n_C \ln\left(\sum_{b=1}^{m} \psi_b \pi_{i_b}\right)$$

and the derivative of $\ln(f_{C\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}))$ with respect to π_i is

$$\frac{\partial \ln(f_{C\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}))}{\partial \pi_i} = \frac{w_{i|C\alpha}}{\pi_i} - \frac{n_C \psi_a}{(\sum_{b=1}^{m} \psi_b \pi_{i_b})},$$

where item T_i appears in position a of the ordered choice set. This derivative will be equal to 0 if item T_i does not appear in the ordered choice set C at all. The derivative of $\ln(f_{C\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}))$ with respect to ψ_a is

$$\frac{\partial \ln(f_{C\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}))}{\partial \psi_a} = \frac{w_{i_a|C\alpha}}{\psi_a} - \frac{n_C \pi_{i_a}}{(\sum_{b=1}^{m} \psi_b \pi_{i_b})},$$

where, once again, item T_{i_a} is the item that appears in position a of the ordered choice set C. We will use these derivatives later to derive an expression for the information matrix for this model. Before we derive the information matrix, we will find an expression for the MLEs for this model.

Since the likelihood function is the product of the density function for a respondent and an ordered choice set over all possible ordered choice sets and over all respondents, we have

$$\begin{aligned} L(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}) &= \prod_{\alpha=1}^{s} \prod_{C} f_{C\alpha}(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}) \\ &= \prod_{C} \frac{\prod_{a=1}^{m} (\psi_a \pi_{i_a})^{w_{i_a|C}}}{(\sum_{b=1}^{m} \psi_b \pi_{i_b})^{s n_C}}, \end{aligned}$$

where $w_{i_a|C} = \sum_{\alpha=1}^{s} w_{i_a|C\alpha}$.

To maximise the likelihood function subject to the constraints of the model, we need to set up a Lagrangian function to incorporate the constraints. For the purposes of convergence, we enforce the normalising constraint present in the MNL model

$$\sum_{i=1}^{t} \ln(\pi_i) = 0.$$

Similarly, we place the constraint

$$\sum_{b=1}^{m} \ln(\psi_b) = 0$$

on the position effects to ensure convergence. We will also constrain the contrasts that are assumed to be negligible. If we let B_a be the matrix containing h such contrasts, then we have

$$B_a \boldsymbol{\gamma} = 0,$$

where, once again, $\boldsymbol{\gamma}$ is the vector containing $\gamma_i = \ln(\pi_i)$. This gives the Lagrangian

$$\begin{aligned} G(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi}) &= \sum_{C} \sum_{a=1}^{m} \left(w_{i_a|C} \ln(\pi_{i_a}) + w_{i_a|C} \ln(\psi_a) - s n_C \ln\Big(\sum_{b=1}^{m} \psi_b \pi_{i_b}\Big)\right) \\ &\quad + \kappa_1 \sum_{i=1}^{t} \ln(\pi_i) + \kappa_2 \sum_{b=1}^{m} \ln(\psi_b) + [\kappa_3 \ldots \kappa_{h+2}] B_a \boldsymbol{\gamma}, \end{aligned}$$

where, once again, there are h contrasts in B_a, and $\kappa_1, \ldots, \kappa_{h+2}$ are Lagrange multipliers. When we differentiate $G(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi})$ with respect to π_i, we obtain

$$\frac{\partial G(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi})}{\partial \pi_i} = \sum_{i \in C} \left(\frac{w_{i|C}}{\pi_i} - \frac{s n_C \psi_a}{(\sum_{b=1}^{m} \psi_b \pi_{i_b})} \right) + \frac{\kappa_1}{\pi_i} + \sum_{x=1}^{h} \kappa_{x+2} (B_a)_{xi} \frac{1}{\pi_i},$$

where item T_i appears in position a of the ordered choice set C. When we differentiate $G(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi})$ with respect to ψ_a, we obtain

$$\frac{\partial G(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi})}{\partial \psi_a} = \sum_{C} \left(\frac{w_{i_a|C}}{\psi_a} - \frac{s n_C \pi_{i_a}}{(\sum_{b=1}^{m} \psi_b \pi_{i_b})} \right) + \frac{\kappa_2}{\psi_a}.$$

As usual, we obtain maximum likelihood estimates by setting these equations equal to 0 and solving simultaneously. This problem can be simplified using matrix notation. Suppose that we let

$$z_i = \sum_{i \in C} w_{i|C} - \frac{s n_C \widehat{\psi_a} \widehat{\pi_i}}{(\sum_{b=1}^{m} \widehat{\psi_b} \widehat{\pi_{i_b}})},$$

where T_i appears in position a of the ordered choice set. By multiplying $\frac{\partial G(\boldsymbol{w}, \boldsymbol{\pi}, \boldsymbol{\psi})}{\partial \pi_i}$ by each π_i in turn, we get

$$z_i + \kappa_1 + \sum_{x=1}^{h} \kappa_{x+2} (B_a)_{xi} = 0,$$

for $i = 1, \ldots, t$. This gives the system of equations

$$\boldsymbol{z} + \kappa_1 \boldsymbol{j}_L^T + B_a^T \boldsymbol{\kappa} = \boldsymbol{0}_L, \tag{5.1}$$

where $\boldsymbol{z} = (z_1, \ldots, z_t)^T$ and $\boldsymbol{\kappa} = (\kappa_3, \ldots, \kappa_{h+2})^T$. Similarly, if we let

$$p_a = \sum_{C} w_{i_a|C} - \frac{s n_C \widehat{\psi_a} \widehat{\pi_{i_a}}}{(\sum_{b=1}^{m} \widehat{\psi_b} \widehat{\pi_{i_b}})},$$

then we obtain

$$\boldsymbol{p} + \kappa_2 \boldsymbol{j}_m^T = \boldsymbol{0}_m$$

as the other set of normal equations.

If we pre–multiply Equation 5.1 by $\boldsymbol{j}_L^T$ we obtain $\kappa_1 = 0$, since $\boldsymbol{j}_L^T \boldsymbol{z}$ is shown to be equal to 0 in Appendix 5.A, and the rows of B_a are the coefficients of contrasts so $\boldsymbol{j}_L^T B_a^T = (B_a \boldsymbol{j}_L)^T = \boldsymbol{0}$. Pre–multiplying Equation 5.1 by B_a, we obtain

$$\boldsymbol{\kappa} = -B_a^T.$$

Substituting this into Equation 5.1, we get

$$(I - B_a^T B_a) \boldsymbol{z} = \boldsymbol{0}_L,$$

and

$$\boldsymbol{p} + \kappa_2 \boldsymbol{j}_m^T = \boldsymbol{0}_m$$

as the normal equations. We obtain the maximum likelihood estimates by solving these equations iteratively.

Now let us look at the estimation of the parameters in the generalised Davidson–Beaver position effects model for our example.

■ **EXAMPLE 5.1.3.**

Recall the experiment considered in Example 5.1.1. Suppose that we present four ordered choice sets to the respondent, as shown in Table 5.2, and are interested in the estimation of the main effects of the attributes and contrasts of the position effects. This means that

$$B_a = \frac{1}{2}\begin{bmatrix} 1 & -1 & -1 & 1 \end{bmatrix},$$

and we are assuming that the contrast for the two–factor interaction

$$\ln(\pi_{00}) - \ln(\pi_{01}) - \ln(\pi_{10}) + \ln(\pi_{11})$$

is zero. Suppose that we present these choice sets to 150 respondents and obtain the set of summarised responses in Table 5.2. Then the Lagrangian for the estimation of the generalised Davidson–Beaver position effects model for this experiment is

$$\begin{aligned} G(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}) &= 54(\ln(\pi_{00})+\ln(\psi_1)) + 10(\ln(\pi_{01})+\ln(\psi_2)) + 86(\ln(\pi_{10})+\ln(\psi_3)) \\ &\quad -150\ln(\psi_1\pi_{00}+\psi_2\pi_{01}+\psi_3\pi_{10}) + 44(\ln(\pi_{01})+\ln(\psi_1)) + 48(\ln(\pi_{00})+\ln(\psi_2)) \\ &\quad +58(\ln(\pi_{11})+\ln(\psi_3)) - 150\ln(\psi_1\pi_{01}+\psi_2\pi_{00}+\psi_3\pi_{11}) + 102(\ln(\pi_{10})+\ln(\psi_1)) \\ &\quad +24(\ln(\pi_{11})+\ln(\psi_2)) + 24(\ln(\pi_{00})+\ln(\psi_3)) - 150\ln(\psi_1\pi_{10}+\psi_2\pi_{11}+\psi_3\pi_{00}) \\ &\quad +55(\ln(\pi_{11})+\ln(\psi_1)) + 82(\ln(\pi_{10})+\ln(\psi_2)) + 13(\ln(\pi_{01})+\ln(\psi_3)) \\ &\quad -150\ln(\psi_1\pi_{11}+\psi_2\pi_{10}+\psi_3\pi_{01}) + \kappa_1(\ln(\pi_{00})+\ln(\pi_{01})+\ln(\pi_{10})+\ln(\pi_{11})) \\ &\quad +\kappa_2(\ln(\psi_1)+\ln(\psi_2)+\ln(\psi_3)) + \kappa_3(\ln(\pi_{00})-\ln(\pi_{01})-\ln(\pi_{10})+\ln(\pi_{11})). \end{aligned}$$

We then differentiate the Lagrangian with respect to each π_i and ψ_a in turn. For example,

$$\begin{aligned} \frac{\partial G(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi})}{\partial \pi_{00}} &= \frac{54+48+24}{\pi_{00}} - \frac{150\psi_1}{\psi_1\pi_{00}+\psi_2\pi_{01}+\psi_3\pi_{10}} - \frac{150\psi_2}{\psi_1\pi_{01}+\psi_2\pi_{00}+\psi_3\pi_{11}} \\ &\quad - \frac{150\psi_3}{\psi_1\pi_{10}+\psi_2\pi_{11}+\psi_3\pi_{00}} + \frac{\kappa_1}{\pi_{00}} + \frac{\kappa_3}{\pi_{00}}, \end{aligned}$$

and

$$\begin{aligned} \frac{\partial G(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi})}{\partial \psi_1} &= \frac{54+44+102+55}{\psi_1} - \frac{150\pi_{00}}{\psi_1\pi_{00}+\psi_2\pi_{01}+\psi_3\pi_{10}} - \frac{150\pi_{01}}{\psi_1\pi_{01}+\psi_2\pi_{00}+\psi_3\pi_{11}} \\ &\quad - \frac{150\pi_{10}}{\psi_1\pi_{10}+\psi_2\pi_{11}+\psi_3\pi_{00}} - \frac{150\pi_{11}}{\psi_1\pi_{11}+\psi_2\pi_{10}+\psi_3\pi_{01}} + \frac{\kappa_2}{\psi_1}. \end{aligned}$$

If we set each of these to 0 and solve iteratively then we obtain the maximum likelihood estimates of the entries in $\boldsymbol{\pi}$ and $\boldsymbol{\psi}$. Since these entries are not estimable without additional constraints, we find contrasts of the entries in $\boldsymbol{\pi}$ and $\boldsymbol{\psi}$. If we let τ_1 be the main effect of the first

Option 1	Option 2	Option 3	$\mathbf{T_1}$	$\mathbf{T_2}$	$\mathbf{T_3}$
0 0	0 1	1 0	54	10	86
0 1	0 0	1 1	44	48	58
1 0	1 1	0 0	102	24	24
1 1	1 0	0 1	55	82	13

Table 5.2: The design and set of responses for the experiment in Example 5.1.3.

attribute, τ_2 the main effect of the second attribute, ψ_L be the linear component of the position effect, and ψ_Q be the quadratic component of the position effect, then we get

$$\widehat{\tau_1} = 0.480 \qquad \widehat{\tau_2} = -0.458 \qquad \widehat{\psi_L} = -0.259 \qquad \widehat{\psi_Q} = 0.121.$$ □

5.2 Properties of the generalised Davidson–Beaver position effects model

In this section, we complete the construction of the information matrix for the estimation of the entries in $\boldsymbol{\gamma}$ and the entries in $\boldsymbol{\psi}$. In practice, not all of these parameters are estimable at the same time so, in the next section, we will transform this information matrix to obtain an information matrix for an estimable set of contrasts of these parameters. Once again, we begin by deriving expressions for the expectations, variances and covariances of the selection indicators, $\boldsymbol{w}$, which were introduced previously. We then use these expressions to simplify the information matrix.

Recall that the entries in $\boldsymbol{w}$ are selection indicators for the choice made by respondent α when presented with an ordered choice set $C = (T_{i_1}, T_{i_2}, \ldots, T_{i_m})$. These w_i have a Bernoulli distribution with expectation

$$\mathcal{E}_\pi(w_{i_a|C,\alpha}) = \frac{\psi_a \pi_{i_a}}{\sum_{b=1}^m \psi_b \pi_{i_b}}.$$

The variance of these selection indicators is given by

$$\text{Var}_\pi(w_{i_a|C,\alpha}) = \frac{\psi_a \pi_{i_a}}{\sum_{b=1}^m \psi_b \pi_{i_b}} \times \frac{\sum_{b=1}^m \psi_b \pi_{i_b} - \psi_a \pi_{i_a}}{\sum_{b=1}^m \psi_b \pi_{i_b}}. \tag{5.2}$$

Next we derive covariances for these selection indicators. As usual, we assume that the selections made in two distinct ordered choice sets are uncorrelated, and thus the w_is of different ordered choice sets have zero correlation. As usual, we also note that for any ordered choice set and any respondent, only one w can be equal to 1, therefore $\mathcal{E}_\pi(w_{i_a|C,\alpha} \times w_{i_{a'}|C,\alpha}) = 0$ for $a \neq b$. Then

$$\begin{aligned}
\text{Cov}_\pi(w_{i_a|C,\alpha}, w_{i_{a'}|C,\alpha}) &= \mathcal{E}_\pi\Big(\big(w_{i_a|C,\alpha} - \mathcal{E}_\pi(w_{i_a|C,\alpha})\big)\big(w_{i_{a'}|C,\alpha} - \mathcal{E}_\pi(w_{i_{a'}|C,\alpha})\big)\Big) \\
&= 0 - \mathcal{E}_\pi(w_{i_a|C,\alpha})\mathcal{E}_\pi(w_{i_{a'}|C,\alpha}) \\
&= -\frac{\psi_a \pi_{i_a}}{\sum_{b=1}^m \psi_b \pi_{i_b}} \times \frac{\psi_{a'} \pi_{i_{a'}}}{\sum_{b=1}^m \psi_b \pi_{i_b}}.
\end{aligned}$$

Then in general,

$$\text{Cov}_\pi(w_{i_a|C,\alpha}, w_{i_{a'}|C',\alpha}) = \begin{cases} -\frac{\psi_a \psi_{a'} \pi_{i_a} \pi_{i_{a'}}}{\sum_{b=1}^m \psi_b \pi_{i_b}}, & \text{if } C = C' \text{ and } a \neq a', \\ \frac{\psi_a \pi_{i_a} (\sum_{b=1}^m \psi_b \pi_{i_b} - \psi_a \pi_{i_a})}{(\sum_{b=1}^m \psi_b \pi_{i_b})^2}, & \text{if } C = C' \text{ and } a = a', \\ 0, & \text{otherwise.} \end{cases} \tag{5.3}$$

We now find the expectations, variances and covariances for the selection indicators in our example.

■ **EXAMPLE 5.2.1.**

Consider the experiment in Example 5.1.1. In particular, consider the first ordered choice set $C = (00, 01, 10)$. The expectation of each of the selection indicators are

$$\begin{aligned}
\mathcal{E}_\pi(w_{00|C,\alpha}) &= \frac{\psi_1 \pi_{00}}{\psi_1 \pi_{00} + \psi_2 \pi_{01} + \psi_3 \pi_{10}}, \\
\mathcal{E}_\pi(w_{01|C,\alpha}) &= \frac{\psi_2 \pi_{01}}{\psi_1 \pi_{00} + \psi_2 \pi_{01} + \psi_3 \pi_{10}},
\end{aligned}$$

and

$$\mathcal{E}_\pi(w_{10|C,\alpha}) = \frac{\psi_3\pi_{10}}{\psi_1\pi_{00}+\psi_2\pi_{01}+\psi_3\pi_{10}}.$$

The variances of the selection indicators are

$$\mathrm{Var}_\pi(w_{00|C,\alpha}) = \frac{\psi_1\pi_{00}(\psi_2\pi_{01}+\psi_3\pi_{10})}{(\psi_1\pi_{00}+\psi_2\pi_{01}+\psi_3\pi_{10})^2},$$

$$\mathrm{Var}_\pi(w_{01|C,\alpha}) = \frac{\psi_2\pi_{01}(\psi_1\pi_{00}+\psi_3\pi_{10})}{(\psi_1\pi_{00}+\psi_2\pi_{01}+\psi_3\pi_{10})^2},$$

and

$$\mathrm{Var}_\pi(w_{10|C,\alpha}) = \frac{\psi_3\pi_{10}(\psi_1\pi_{00}+\psi_2\pi_{01})}{(\psi_1\pi_{00}+\psi_2\pi_{01}+\psi_3\pi_{10})^2}.$$

The covariances of pairs of these selection indicators are

$$\mathrm{Cov}_\pi(w_{00|C,\alpha}, w_{01|C,\alpha}) = \frac{-\psi_1\psi_2\pi_{00}\pi_{01}}{(\psi_1\pi_{00}+\psi_2\pi_{01}+\psi_3\pi_{10})^2},$$

$$\mathrm{Cov}_\pi(w_{00|C,\alpha}, w_{10|C,\alpha}) = \frac{-\psi_1\psi_3\pi_{00}\pi_{10}}{(\psi_1\pi_{00}+\psi_2\pi_{01}+\psi_3\pi_{10})^2},$$

and

$$\mathrm{Cov}_\pi(w_{01|C,\alpha}, w_{10|C,\alpha}) = \frac{-\psi_2\psi_3\pi_{01}\pi_{10}}{(\psi_1\pi_{00}+\psi_2\pi_{01}+\psi_3\pi_{10})^2}.$$

□

We now construct the information matrix for the estimation of the generalised Davidson–Beaver position effects model. As with the Davidson–Beaver position effects model, the construction will be easier if we partition the information matrix into four blocks. $I_{\pi\pi}(\boldsymbol{\pi},\boldsymbol{\psi})$ contains minus the expected values of the second derivatives of the density function with respect to two of the entries in $\boldsymbol{\pi}$. $I_{\psi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})$ contains minus the expected values of the second derivatives of the density function with respect to two entries in $\boldsymbol{\psi}$. $I_{\pi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})$ and $I_{\psi\pi}(\boldsymbol{\pi},\boldsymbol{\psi})$ contain minus the expected value of the second derivatives with respect to one entry in $\boldsymbol{\pi}$ and one entry in $\boldsymbol{\psi}$. The partitioned matrix is denoted by

$$I(\boldsymbol{\pi},\boldsymbol{\psi}) = \begin{bmatrix} I_{\pi\pi}(\boldsymbol{\pi},\boldsymbol{\psi}) & I_{\pi\psi}(\boldsymbol{\pi},\boldsymbol{\psi}) \\ I_{\psi\pi}(\boldsymbol{\pi},\boldsymbol{\psi}) & I_{\psi\psi}(\boldsymbol{\pi},\boldsymbol{\psi}) \end{bmatrix},$$

where $I_{\pi\psi}(\boldsymbol{\pi},\boldsymbol{\psi}) = (I_{\psi\pi}(\boldsymbol{\pi},\boldsymbol{\psi}))^T$.

El-Helbawy and Bradley [1978] states that, under some mild regularity conditions, as given in Section 1.1, the $(i,j)^{\text{th}}$ entry of the information matrix for a discrete choice experiment without position effects is

$$I(\boldsymbol{\pi})_{ij} = \sum_{q=1}^{t-1}\sum_{r=q+1}^{t} \frac{n_{qr}}{N}\mathcal{E}_\pi\left(\left(\frac{\partial \ln(f_{qr\alpha}(\boldsymbol{\pi},\boldsymbol{w}))}{\partial \pi_i}\right)\left(\frac{\partial \ln(f_{qr\alpha}(\boldsymbol{\pi},\boldsymbol{w}))}{\partial \pi_j}\right)\right).$$

Then we can use the derivatives found earlier, as well as the variance and covariance expressions given above, to simplify the information matrix. Since $m \geq 2$ in this case, we sum over the choice sets of size m rather than pairs of items, and modify the notation for n_{qr} and $f_{qr\alpha}(\boldsymbol{\pi},\boldsymbol{w})$ accordingly. Again, it is convenient to look at one block matrix at a time, and the diagonal entries of a block matrix separately to the off–diagonal entries.

We begin with a generic off–diagonal entry of $I_{\pi\pi}(\boldsymbol{\pi},\boldsymbol{\psi})$, $I_{\pi\pi}(\boldsymbol{\pi},\boldsymbol{\psi})_{ij}$. We have

$$I_{\pi\pi}(\boldsymbol{\pi},\boldsymbol{\psi})_{ij} = \sum_C \frac{n_C}{N}\mathcal{E}_\pi\Big(\Big(\frac{\partial \ln(f_{C\alpha}(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}))}{\partial \pi_i}\Big)\Big(\frac{\partial \ln(f_{C\alpha}(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}))}{\partial \pi_j}\Big)\Big).$$

Since the derivative of $\ln(f_{C\alpha}(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}))$ with respect to π_i will be 0 unless the item T_i appears in choice set C, we can restrict our summation to those ordered choice sets that include both T_i and T_j. This gives

$$I_{\pi\pi}(\boldsymbol{\pi},\boldsymbol{\psi})_{ij} = \sum_{C|T_i,T_j\in C} \frac{n_C}{N}\mathcal{E}_\pi\Big(\Big(\frac{\partial \ln(f_{C\alpha}(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}))}{\partial \pi_i}\Big)\Big(\frac{\partial \ln(f_{C\alpha}(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}))}{\partial \pi_j}\Big)\Big).$$

Straight–forward differentiation gives

$$\frac{\partial \ln(f_{C\alpha}(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}))}{\partial \pi_i} = \frac{w_{i|C,\alpha}}{\pi_i} - \mathcal{E}_\pi\Big(\frac{w_{i|C,\alpha}}{\pi_i}\Big). \tag{5.4}$$

Thus we get

$$I_{\pi\pi}(\boldsymbol{\pi},\boldsymbol{\psi})_{ij} = \sum_{C|T_i,T_j\in C} \frac{n_C}{N}\mathrm{Cov}_\pi\Big(\frac{w_{i|C,\alpha}}{\pi_i},\frac{w_{j|C,\alpha}}{\pi_j}\Big).$$

If we then substitute the covariance, given in Equation 5.3, we obtain

$$I_{\pi\pi}(\boldsymbol{\pi},\boldsymbol{\psi})_{ij} = \sum_{C|T_i,T_j\in C} \frac{n_C}{N}\frac{-\psi_{a_i}\psi_{a_j}}{(\sum_{b=1}^m \psi_b\pi_{i_b})^2},$$

where item T_x appears in position a_x of the ordered choice set, for all x.

Now let us consider a generic diagonal term in $I_{\pi\pi}(\boldsymbol{\pi},\boldsymbol{\psi})$. Again, noting that the derivative of $\ln(f_{C\alpha}(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}))$ with respect to π_i is 0 when item T_i is not in the ordered choice set C, we have

$$I_{\pi\pi}(\boldsymbol{\pi},\boldsymbol{\psi})_{ii} = \sum_{C|T_i\in C} \frac{n_C}{N}\mathcal{E}_\pi\Big(\Big(\frac{\partial \ln(f_{C\alpha}(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}))}{\partial \pi_i}\Big)^2\Big).$$

Using the result in Equation 5.4, this becomes

$$I_{\pi\pi}(\boldsymbol{\pi},\boldsymbol{\psi})_{ii} = \sum_{C|T_i\in C} \frac{n_C}{N}\mathrm{Var}_\pi\Big(\frac{w_{i|C,\alpha}}{\pi_i}\Big).$$

When we substitute the variance, given in Equation 5.2, we obtain

$$I_{\pi\pi}(\boldsymbol{\pi},\boldsymbol{\psi})_{ii} = \sum_{C|T_i\in C} \frac{n_C}{N\pi_i}\frac{\psi_{a_i}(\sum_{b=1}^m \psi_b\pi_{i_b} - \psi_{a_i}\pi_i)}{(\sum_{b=1}^m \psi_b\pi_{i_b})^2}.$$

We now repeat this process to evaluate a generic entry in $I_{\pi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})$. Once again, we exclude choice sets that do not include the item T_i, and so

$$I_{\pi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})_{ia} = \sum_{C|T_i\in C} \frac{n_C}{N}\mathcal{E}_\pi\Big(\Big(\frac{\partial \ln(f_{C\alpha}(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}))}{\partial \pi_i}\Big)\Big(\frac{\partial \ln(f_{C\alpha}(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}))}{\partial \psi_a}\Big)\Big).$$

Straight–forward differentiation of the density function for a choice set and respondent with respect to ψ_a gives

$$\frac{\partial \ln(f_{C\alpha}(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}))}{\partial \psi_a} = \frac{w_{i_a|C,\alpha}}{\psi_a} - \mathcal{E}_\pi\Big(\frac{w_{i_a|C,\alpha}}{\psi_a}\Big). \tag{5.5}$$

This result, together with Equation 5.4, gives

$$\begin{aligned} I_{\pi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})_{ia} &= \sum_{C|T_i\in C} \frac{n_C}{N}\text{Cov}_\pi\Big(\frac{w_{i|C,\alpha}}{\pi_i}, \frac{w_{i_a|C,\alpha}}{\psi_a}\Big) \\ &= \sum_{C|T_i\in C} \frac{n_C\left(\delta_{T_i \text{ is in pos } a} \times \text{Var}_\pi(w_{i|C,\alpha}) + (1-\delta_{T_i \text{ is in pos } a})\text{Cov}_\pi(w_{i|C,\alpha}, w_{i_a|C,\alpha})\right)}{N\psi_a\pi_i}, \end{aligned}$$

where $\delta_{T_i \text{ is in pos } a}$ is an indicator that equals 1 if T_i appears in position a of the ordered choice set C, and 0 otherwise. We can then substitute the variance and covariance expressions, given in Equations 5.2 and 5.3, to obtain

$$I_{\pi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})_{ia} = \sum_{C|T_i\in C} \frac{n_C}{N}\Bigg(\frac{\delta_{T_i \text{ is in pos } a}(\sum_{b=1}^m \psi_b\pi_{i_b} - \psi_a\pi_i)}{(\sum_{b=1}^m \psi_b\pi_{i_b})^2} - \frac{(1-\delta_{T_i \text{ is in pos } a})\psi_{a_i}\pi_{i_a}}{(\sum_{b=1}^m \psi_b\pi_{i_b})^2}\Bigg).$$

Finally, we look at $I_{\psi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})$. We begin with the off–diagonal entries in each block,

$$I_{\psi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})_{a_1a_2} = \sum_C \frac{n_C}{N}\mathcal{E}_\pi\Bigg(\Big(\frac{\partial \ln(f_{C\alpha}(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}))}{\partial\psi_{a_1}}\Big)\Big(\frac{\partial \ln(f_{C\alpha}(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}))}{\partial\psi_{a_2}}\Big)\Bigg),$$

where $a_1 \neq a_2$.

In this case, there are no derivatives with respect to any of the π_is, therefore no choice sets need to be excluded from the summation. We can use Equation 5.5 to obtain

$$I_{\psi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})_{a_1a_2} = \sum_C \frac{n_C}{N}\text{Cov}_\pi\Big(\frac{w_{i_{a_1}|C,\alpha}}{\psi_{a_1}}, \frac{w_{i_{a_2}|C,\alpha}}{\psi_{a_2}}\Big).$$

Since $a_1 \neq a_2$, we substitute the covariance of the two selection indicators, giving

$$I_{\psi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})_{a_1a_2} = \sum_C \frac{n_C}{N\psi_{a_1}\psi_{a_2}} \frac{-\pi_{i_{a_1}}\pi_{i_{a_2}}}{(\sum_{b=1}^m \psi_b\pi_{i_b})^2}.$$

Finally, we consider the diagonal entries of $I_{\psi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})$. We begin with

$$I_{\psi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})_{aa} = \sum_C \frac{n_C}{N}\mathcal{E}_\pi\Bigg(\Big(\frac{\partial \ln(f_{C\alpha}(\boldsymbol{w},\boldsymbol{\pi},\boldsymbol{\psi}))}{\partial\psi_a}\Big)^2\Bigg).$$

Using Equation 5.5, this can be simplified to give

$$I_{\psi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})_{aa} = \sum_C \frac{n_C}{N}\text{Var}_\pi\Big(\frac{w_{i_a|C,\alpha}}{\psi_a}\Big),$$

which, when we substitute the variance given in Equation 5.2, becomes

$$I_{\psi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})_{aa} = \sum_C \frac{n_C}{N\psi_a} \frac{\pi_{i_a}((\sum_{b=1}^m \psi_b\pi_{i_b}) - \psi_a\pi_{i_a})}{(\sum_{b=1}^m \psi_b\pi_{i_b})^2},$$

where the summation is over all choice sets C.

Once again, our ultimate goal is to estimate contrasts of the entries in $\boldsymbol{\gamma} = \ln(\boldsymbol{\pi})$ and contrasts of the entries in $\boldsymbol{\psi}$, and not the entries in $\boldsymbol{\pi}$ and $\boldsymbol{\psi}$ themselves, since they are not estimable without additional constraints. One way to do this is to first find the information matrix for the estimation of the entries in $\boldsymbol{\gamma}$ and $\boldsymbol{\psi}$. The equivalent matrix for the MNL model was introduced in Section 1.1. We partition this matrix in the same way as in Section 4.2 to give

$$\Lambda(\boldsymbol{\pi},\boldsymbol{\psi}) = \begin{bmatrix} \Lambda_{\gamma\gamma}(\boldsymbol{\pi},\boldsymbol{\psi}) & \Lambda_{\gamma\psi}(\boldsymbol{\pi},\boldsymbol{\psi}) \\ \Lambda_{\psi\gamma}(\boldsymbol{\pi},\boldsymbol{\psi}) & \Lambda_{\psi\psi}(\boldsymbol{\pi},\boldsymbol{\psi}) \end{bmatrix}.$$

As was the case when estimating the Davidson–Beaver position effects model for $m = 2$, we have

$$\Lambda(\boldsymbol{\pi}, \boldsymbol{\psi}) = PI(\boldsymbol{\pi}, \boldsymbol{\psi})P^T,$$

where

$$P = \begin{bmatrix} \pi_1 & 0 & \dots & 0 & \mathbf{0} \\ 0 & \pi_2 & \dots & 0 & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \dots & \pi_t & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & I_m \end{bmatrix}.$$

If we apply this transformation to each of the generic entries, we get

$$\begin{aligned} \Lambda_{\gamma\gamma}(\boldsymbol{\pi}, \boldsymbol{\psi})_{ij} &= \pi_i \pi_j I_{\pi\pi}(\boldsymbol{\pi}, \boldsymbol{\psi})_{ij}, \\ \Lambda_{\gamma\gamma}(\boldsymbol{\pi}, \boldsymbol{\psi})_{ii} &= \pi_i^2 I_{\pi\pi}(\boldsymbol{\pi}, \boldsymbol{\psi})_{ii}, \\ \Lambda_{\gamma\psi}(\boldsymbol{\pi}, \boldsymbol{\psi})_{ia} &= \pi_i I_{\pi\psi}(\boldsymbol{\pi}, \boldsymbol{\psi})_{ia}, \end{aligned}$$

and

$$\Lambda_{\psi\psi}(\boldsymbol{\pi}, \boldsymbol{\psi}) = I_{\psi\psi}(\boldsymbol{\pi}, \boldsymbol{\psi}).$$

If we assume, as did Davidson and Beaver [1977], the null hypothesis of equal merits for each of the items, and that the entries of $\boldsymbol{\psi}$ are left unspecified, then the expressions for the entries of $\Lambda(\boldsymbol{\pi}, \boldsymbol{\psi})$ simplify greatly. If we let

$$\boldsymbol{\pi} = \boldsymbol{j}_t = \boldsymbol{\pi}_0,$$

then we obtain

$$\begin{aligned} \Lambda_{\pi\pi}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{ij} &= \sum_{C|T_i,T_j \in C} \frac{n_C}{N} \frac{(-\psi_{a_i}\psi_{a_j})}{\Psi_1} \\ &= -\frac{1}{\Psi_1} \sum_{a=1}^{m} \sum_{b \neq a} \psi_a \psi_b \lambda_{T_i \text{ in pos } a, T_j \text{ in pos } b}, \end{aligned}$$

where

$$\Psi_1 = \left(\sum_{b=1}^{m} \psi_b \right)^2$$

and

$$\lambda_{T_i \text{ in pos } a, T_j \text{ in pos } b} = \sum_{C|T_i,T_j \in C} \frac{n_C}{N} \delta_{T_i \text{ is in pos } a} \times \delta_{T_j \text{ is in pos } b}.$$

A similar argument gives

$$\begin{aligned} \Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{ii} &= \sum_{C|T_i \in C} \frac{n_C}{N} \frac{\psi_{a_i}(\sum_{b=1}^{m} \psi_b - \psi_{a_i})}{\Psi_1} \\ &= \frac{1}{\Psi_1} \sum_{a=1}^{m} \psi_a \Big(\sum_{b=1}^{m} \psi_b - \psi_a \Big) \lambda_{T_i \text{ in pos } a}, \end{aligned}$$

where

$$\lambda_{T_i \text{ in pos } a} = \sum_{C|T_i \in C} \frac{n_C}{N} \delta_{T_i \text{ is in pos } a}.$$

The generic term in $\Lambda_{\pi\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi})$ becomes

$$\begin{aligned}\Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{ia} &= \sum_{C|T_i \in C} \frac{n_C}{N}\left(\delta_{T_i \text{ is in pos } a} \times \frac{\sum_{b=1}^m \psi_b - \psi_a}{\Psi_1} - (1-\delta_{T_i \text{ is in pos } a}) \times \frac{\psi_{a_i}\pi_{i_a}}{\Psi_1}\right) \\ &= \frac{1}{\Psi_1}\sum_{b \neq a} \psi_b(\lambda_{T_i \text{ in pos } a} - \lambda_{T_i \text{ in pos } b}).\end{aligned}$$

The off-diagonal entries of $\Lambda_{\psi\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi})$ become

$$\begin{aligned}\Lambda_{\psi\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{a_1a_2} &= \sum_C \frac{-n_C}{N\Psi_1} \\ &= -\frac{1}{\Psi_1},\end{aligned}$$

since $\sum_C n_C = N$. Finally, the diagonal entries of $\Lambda_{\psi\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi})$ become

$$\begin{aligned}\Lambda_{\psi\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{aa} &= \sum_C \frac{n_C}{N\psi_a}\frac{(\sum_{b=1}^m \psi_b) - \psi_a}{\Psi_1} \\ &= \frac{(\sum_{b=1}^m \psi_b) - \psi_a}{\psi_a\Psi_1}.\end{aligned}$$

We notice that the entries in $\Lambda_{\psi\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi})$ depend only on the entries in $\boldsymbol{\psi}$ under the null hypothesis. Therefore this block matrix is independent of the design, given a fixed choice set size.

Now let us find the $\Lambda(\boldsymbol{\pi}_0, \boldsymbol{\psi})$ matrix in our example.

■ **EXAMPLE 5.2.2.**
Recall the experiment introduced in Example 5.1.1. The information matrix for the estimation of the $\boldsymbol{\gamma}$ and $\boldsymbol{\psi}$ under the null hypothesis of equal merits is shown in Table 5.3. □

5.3 Representing options using k attributes

In this section we consider the construction of the information matrix for the estimation of contrasts of the entries in $\boldsymbol{\gamma}$ and contrasts of the entries in $\boldsymbol{\psi}$. In particular, we are interested in the estimation of the contrasts of the entries in $\boldsymbol{\gamma}$ that relate to the main effects and interaction effects of the attributes as introduced in Chapter 1.

We begin by constructing a matrix B that contains the coefficients of the contrasts of the entries in $\boldsymbol{\gamma}$ and the contrasts of the entries in $\boldsymbol{\psi}$. We will assume that any interactions between the attributes in the experiment and position are not of interest. We can then partition the B matrix to give

$$B = \begin{bmatrix} B_\gamma & \mathbf{0} \\ \mathbf{0} & B_\psi \end{bmatrix}.$$

Then the information matrix for the estimation of the contrasts $B_\gamma\boldsymbol{\gamma}$ and $B_\psi\boldsymbol{\psi}$ under the null hypothesis of equal merits and making no assumption about the magnitude of the position effects is given by

$$\begin{aligned}C(\boldsymbol{\pi}_0, \boldsymbol{\psi}) &= B\Lambda(\boldsymbol{\pi}_0, \boldsymbol{\psi})B^T \\ &= \begin{bmatrix} B_\gamma\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})B_\gamma^T & B_\gamma\Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi})B_\psi^T \\ B_\psi\Lambda_{\psi\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})B_\gamma^T & B_\psi\Lambda_{\psi\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi})B_\psi^T \end{bmatrix}.\end{aligned}$$

Let us apply these results to the experiment in Example 5.1.1.

$$\Lambda(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = \frac{1}{2\left(\psi_1+\psi_2+\psi_3\right)^2} \times$$

$$\begin{bmatrix}
\psi_2\psi_3+\psi_1\left(\psi_2+\psi_3\right) & -\psi_1\psi_2 & -\psi_1\psi_3 & -\psi_2\psi_3 & 0 & 0 & 0 \\
-\psi_1\psi_2 & \psi_2\psi_3+\psi_1\left(\psi_2+\psi_3\right) & -\psi_2\psi_3 & -\psi_1\psi_3 & 0 & 0 & 0 \\
-\psi_1\psi_3 & -\psi_2\psi_3 & \psi_2\psi_3+\psi_1\left(\psi_2+\psi_3\right) & -\psi_1\psi_2 & 0 & 0 & 0 \\
-\psi_2\psi_3 & -\psi_1\psi_3 & -\psi_1\psi_2 & \psi_2\psi_3+\psi_1\left(\psi_2+\psi_3\right) & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{2(\psi_2+\psi_3)}{\psi_1} & -2 & -2 \\
0 & 0 & 0 & 0 & -2 & \frac{2(\psi_1+\psi_3)}{\psi_2} & -2 \\
0 & 0 & 0 & 0 & -2 & -2 & \frac{2(\psi_1+\psi_2)}{\psi_3}
\end{bmatrix}.$$

Table 5.3: The information matrix for the estimation of $\boldsymbol{\gamma}$ and $\boldsymbol{\psi}$ in Example 5.2.2

■ **EXAMPLE 5.3.1.**
Consider the experiment introduced in Example 5.1.1 and the design introduced in Example 5.1.3 for the estimation of the main effects of the attributes and the linear and quadratic components of the position effect. The contrast matrix for the estimation of the main effects of the attributes and contrasts of the position effects is

$$B = \begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \end{bmatrix}.$$

We see that B_γ is a 2×4 matrix of contrast coefficients and B_ψ a 2×3 matrix of contrast coefficients. Then the information matrix for the estimation of the main effects of the attributes and contrasts of the position effects is

$$C(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = \begin{bmatrix} \frac{(\psi_1+\psi_2)\psi_3}{(\psi_1+\psi_2+\psi_3)^2} & 0 & 0 & 0 \\ 0 & \frac{\psi_2(\psi_1+\psi_3)}{(\psi_1+\psi_2+\psi_3)^2} & 0 & 0 \\ 0 & 0 & \frac{\psi_1+\psi_3}{2\psi_1\psi_3(\psi_1+\psi_2+\psi_3)} & \frac{\psi_1-\psi_3}{2\sqrt{3}\psi_1\psi_3(\psi_1+\psi_2+\psi_3)} \\ 0 & 0 & \frac{\psi_1-\psi_3}{2\sqrt{3}\psi_1\psi_3(\psi_1+\psi_2+\psi_3)} & \frac{\psi_2\psi_3+\psi_1(\psi_2+4\psi_3)}{6\psi_1\psi_2\psi_3(\psi_1+\psi_2+\psi_3)} \end{bmatrix}.$$

We are able to estimate the main effects of the attributes and the position effect contrasts independently when using this design, since the information matrix is block diagonal. □

Now that we have the information matrix for the estimation of the effects that are of interest, we will next find results about the optimal designs for the estimation of these effects.

5.4 Optimal designs for the generalised Davidson–Beaver position effects model

In this section, we determine the optimal designs for the estimation of the main effects of the attributes and contrasts of the position effects when the generalised Davidson–Beaver position effects model is used. We will begin by finding design constraints that allow the information matrix to be block diagonal, and then determine optimal designs for the estimation of the main effects of the attributes and contrasts of the position effects. We do this by showing that, when the information matrix for the estimation of the main effects of the attributes and contrasts of the position effects is block diagonal with respect to the attributes, we can consider each attribute separately when finding the optimal set of generators used to construct the choice design from the starting design. We conclude by looking at the form of the determinant of the information matrix when the main effects plus two–factor interactions of the attributes and specific contrasts of position effects are of interest.

We begin with the estimation of the main effects of the attributes and contrasts of the position effects.

5.4.1 Main Effects plus Position Effects

In this subsection, we consider optimal designs for the estimation of the main effects of the attributes and contrasts of the position effects. In the next section, we extend this to the estimation of the main effects plus two–factor interactions of the attributes and contrasts of the position effects.

It is desirable to be able to estimate the attribute effects independently of the position effects. This requires $C_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = \mathbf{0}$, and therefore $C_{\psi\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = \mathbf{0}$, since $C_{\psi\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = C_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi})^T$. Once again, we need an expression for the entries in this block matrix before we can find the conditions for a block diagonal information matrix.

Recall that the $(1, 2)$ block of $C(\boldsymbol{\pi}_0, \boldsymbol{\psi})$ is equal to

$$C_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = B_\gamma \Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) B_\psi^T.$$

A generic entry of the matrix obtained by multiplying the first two of these matrices together is

$$\begin{aligned}(B_\gamma \Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}))_{ia} &= \sum_{j=1}^{t} B_{ij} \frac{1}{\Psi_1} \sum_{b \neq a} \psi_b (\lambda_{T_i \text{ in pos } a} - \lambda_{T_i \text{ in pos } b}) \\ &= \frac{1}{\Psi_1} \sum_{b \neq a} \psi_b \sum_{j=1}^{t} B_{ij} (\lambda_{T_i \text{ in pos } a} - \lambda_{T_i \text{ in pos } b}).\end{aligned}$$

While it is obvious that $C_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = \mathbf{0}$ if all of the items that appear in the experiment appear in each position of the choice set equally often, this constraint on the design is far more restrictive than necessary. This constraint would allow us to estimate all of the higher order effects independently of the contrasts of the position effects. Usually, we are only interested in the estimation of the main effects of attributes, or perhaps the estimation of the main effects plus two–factor interactions of the attributes, and not all estimable contrasts of the entries in $\boldsymbol{\gamma}$. Therefore we only need the main effects of the attributes to be independent of the position effects.

We now prove a lemma that gives conditions for $C_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = \mathbf{0}$ when only the main effects of the attributes and contrasts of the position effects are of interest.

■ **LEMMA 5.4.1.**
The information matrix for the estimation of the main effects of the attributes and contrasts of the position effects is block diagonal with respect to the main effects of the attributes and the contrasts of the position effects if each of the levels for each attribute appear in each position equally often.

Proof. Items that have the attribute of interest at the same level will have the same contrast coefficient. Then for the j^{th} contrast, corresponding to a component of the main effect of attribute q,

$$(B_\gamma \Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}))_{ja} = \frac{1}{\Psi_1} \sum_{b \neq a} \psi_b \sum_{x=1}^{\ell_q} B_{jx} (\lambda_{\text{att } q=x \text{ in pos } a} - \lambda_{\text{att } q=x \text{ in pos } b}).$$

Then $(B_\gamma \Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}))_{ja} = 0$ if $\lambda_{\text{att } q=x \text{ in pos } a} - \lambda_{\text{att } q=x \text{ in pos } b} = 0$ for all attribute levels, $0 \leq x \leq \ell_q - 1$, and all positions, $1 \leq b \leq m$ where $a \neq b$. It follows that $B_\gamma \Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = \mathbf{0}$ if, for each attribute, each attribute level appears in each position in the choice set equally often. If this is the case, then the information matrix for the estimation of the main effects of the attributes and contrasts of the position effects will be block diagonal, with one block corresponding to the main effects of the attributes and one block corresponding to the position effects. □

We now develop theory on the structure of the optimal design of choice experiments when the generalised Davidson–Beaver position effects model is used. In order to prove a result similar to that in Theorem 1.3.6 for the estimation of the main effects of the attributes and contrasts of the position effects, we need to modify some of the definitions introduced in Section 1.1 to accommodate the importance of order. Let

- $c_{\boldsymbol{v}_j,a}$ be the number of choice sets with ordered difference vector $\boldsymbol{v}_j$ that contain the item $00\ldots0$ in position a of the choice set,
- $x_{\boldsymbol{v}_j;\boldsymbol{d},a,b}$ be the number of times that the difference $\boldsymbol{d}$ appears as the difference between positions a and b in ordered difference vector $\boldsymbol{v}_j$ (i.e. $T_{i_a}+\boldsymbol{d}=T_{i_b}$). Note that $\sum_{\boldsymbol{d}} x_{\boldsymbol{v}_j;\boldsymbol{d},a,b} = 1$ and $\sum_{a\neq b}\sum_{\boldsymbol{d}} x_{\boldsymbol{v}_j;\boldsymbol{d},a,b} = \binom{m}{2}$,
- $i_{\boldsymbol{v}_j}$ be an indicator variable that equals 1 if all of the choice sets with an ordered difference vector $\boldsymbol{v}_j$ appear in the choice experiment, and 0 if none of the choice sets with ordered difference vector $\boldsymbol{v}_j$ appear in the experiment.

We now illustrate these constants with an example.

■ **EXAMPLE 5.4.1.**
Recall the experiment considered in Example 5.1.1 with two 2–level attributes presented in choice sets of size 3. There are 6 possible ordered difference vectors, which are shown in Table 5.4. The first entry in each difference vector is the difference between the first and second items in the choice set, the second entry is the difference between the first and third items in the choice set, and the third entry is the difference between the second and third items in the choice set.

$\boldsymbol{v}_1$	$(01, 10, 11)$
$\boldsymbol{v}_2$	$(01, 11, 10)$
$\boldsymbol{v}_3$	$(10, 01, 11)$
$\boldsymbol{v}_4$	$(10, 11, 01)$
$\boldsymbol{v}_5$	$(11, 01, 10)$
$\boldsymbol{v}_6$	$(11, 10, 01)$

Table 5.4: Possible ordered difference vectors for the experiment in Example 5.4.1.

The experiment in Table 5.2 contains all of the choice sets with ordered difference vector $\boldsymbol{v}_1$. Therefore $i_{\boldsymbol{v}_1}=1$, and $i_{\boldsymbol{v}_j}=0$ for all of the other difference vectors. The item 00 appears in each position once, thus

$$c_{\boldsymbol{v}_1,1} = c_{\boldsymbol{v}_1,2} = c_{\boldsymbol{v}_1,3} = 1.$$

Within $\boldsymbol{v}_1$, we have

$$x_{\boldsymbol{v}_1;(01),1,2} = x_{\boldsymbol{v}_1;(10),1,3} = x_{\boldsymbol{v}_1;(11),2,3} = 1,$$

and all other $x_{\boldsymbol{v}_1;\boldsymbol{d},a,b} = 0$. □

Using these definitions, we can give a general form for $\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})$ when the generalised Davidson–Beaver position effects model is used.

■ **THEOREM 5.4.2.**

Under the usual null hypothesis of equal merits

$$\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = \frac{\Psi_2}{\Psi_1} z I_L - \frac{1}{\Psi_1} \sum_{\boldsymbol{d}} \sum_{a=1}^{m} \sum_{b \neq a} \psi_a \psi_b y_{\boldsymbol{d},a,b} D_{\boldsymbol{d}},$$

where

$$y_{\boldsymbol{d},a,b} = \frac{1}{N \prod_{q=1}^{m} (\ell_q - 1)^{i_q}} \sum_j c_{\boldsymbol{v}_j,1} i_{\boldsymbol{v}_j} x_{\boldsymbol{v}_j;\boldsymbol{d},a,b},$$

$$\Psi_2 = \sum_{a \neq b} \psi_a \psi_b, \qquad \text{and } z = \frac{1}{N} \sum_j c_{\boldsymbol{v}_j,a} i_{\boldsymbol{v}_j}.$$

The summations over j and $\boldsymbol{d}$ are over all possible difference vectors $\boldsymbol{v}_j$ and all distinct difference vector entries $\boldsymbol{d}$ respectively. □

Proof. We showed earlier that under the null hypothesis of equal merits, the diagonal elements of $\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})$ are given by

$$\begin{aligned} \Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{ii} &= \frac{1}{\Psi_1} \sum_{a=1}^{m} \left[\lambda_{T_i \text{ in pos } a} \times \psi_a \left(\sum_{b=1}^{m} \psi_b - \psi_a \right) \right] \\ &= \frac{1}{\Psi_1} \sum_{a=1}^{m} \left[\frac{\delta_{T_i \text{ in pos } a}}{N} \times \psi_a \left(\sum_{b=1}^{m} \psi_b - \psi_a \right) \right]. \end{aligned}$$

In the class of competing designs discussed in Section 1.1 we assume that all choice sets with ordered difference vector $\boldsymbol{v}_j$ will appear in the experiment if $i_{\boldsymbol{v}_j} = 1$. It follows that

$$c_{\boldsymbol{v}_j,1} = c_{\boldsymbol{v}_j,2} = \ldots = c_{\boldsymbol{v}_j,m} = c_{\boldsymbol{v}_j},$$

and that $\lambda_{T_i \text{ in pos } a} = \frac{1}{N} \sum_{a=1}^{m} \sum_j c_{\boldsymbol{v}_j} i_{\boldsymbol{v}_j}$. Therefore

$$\begin{aligned} \Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{ii} &= \frac{1}{N\Psi_1} \sum_j \sum_{a=1}^{m} c_{\boldsymbol{v}_j} i_{\boldsymbol{v}_j} \psi_a \left(\sum_{b=1}^{m} \psi_b - \psi_a \right) \\ &= \frac{\Psi_2}{\Psi_1} \times z. \end{aligned}$$

Under the null hypotheses of equal merits, the off–diagonal elements of the $\Lambda_{\gamma\gamma}(\boldsymbol{\pi}, \boldsymbol{\psi})$ matrix are

$$\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{ij} = -\frac{1}{\Psi_1} \sum_{a=1}^{m} \sum_{b \neq a} \psi_a \psi_b \lambda_{T_i \text{ in pos } a, T_j \text{ in pos } b}.$$

We need to find an expression for the proportion of choice sets with item T_i in position a of the choice set and item T_j in position b of the choice set in terms of which difference vectors are used in the experiment. To do this, we enumerate the number of choice sets that have this pair of items in this pair of positions.

There are $Lc_{\boldsymbol{v}_j}$ possible choice sets with difference vector $\boldsymbol{v}_j$. By the definition of the class of competing designs, there are $Lc_{\boldsymbol{v}_j} i_{\boldsymbol{v}_j}$ choice sets with difference vector $\boldsymbol{v}_j$ in the experiment. Then there are

$$\sum_j Lc_{\boldsymbol{v}_j} i_{\boldsymbol{v}_j}$$

choice sets in the experiment in total. It follows that the number of these choice sets with difference $\boldsymbol{d}$ between positions a and b of the choice set is

$$\sum_j Lc_{\boldsymbol{v}_j} i_{\boldsymbol{v}_j} x_{\boldsymbol{v}_j;\boldsymbol{d},a,b}.$$

If the pair T_i and T_j have a difference $\boldsymbol{d}$, then we find that only one pair of items with $\boldsymbol{d}$ will be the pair T_i and T_j. Then we need to enumerate the number of pairs of items with difference $\boldsymbol{d}$.

The number of pairs with difference $\boldsymbol{d}$ depends on which entries in the difference are 0, and which are 1. If $i_q = 0$ for an attribute q, then the level of the q^{th} attribute must be the same. If $i_q = 1$, then there are $\ell_q - 1$ possible levels for the q^{th} attribute that would allow a difference of $\boldsymbol{d}$ from T_i. Then the number of items with difference $\boldsymbol{d}$ from T_i is

$$\Gamma_{\boldsymbol{d}} = \prod_{q=1}^{k} (\ell_q - 1)^{i_q}.$$

If items T_i and T_j have a difference $\boldsymbol{d}$, the proportion of choice sets in the experiment that contain T_i in position a and T_j in position b is

$$y_{\boldsymbol{d},a,b} = \frac{1}{N\Gamma_{\boldsymbol{d}}} \sum_j c_{\boldsymbol{v}_j} i_{\boldsymbol{v}_j} x_{\boldsymbol{v}_j;\boldsymbol{d},a,b}.$$

The off–diagonal elements of $\Lambda_{\gamma\gamma}(\boldsymbol{\pi}, \boldsymbol{\psi})$ can then be expressed as

$$\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = -\frac{1}{\Psi_1} \sum_{a=1}^{m} \sum_{b \neq a} \psi_a \psi_b \sum_{\boldsymbol{d}} y_{\boldsymbol{d},a,b} D_{\boldsymbol{d}},$$

where $D_{\boldsymbol{d}}$ is a $t \times t$ matrix with entries either 0 or 1 such that there is a 1 in position (i,j) if and only if items T_i and T_j have a difference $\boldsymbol{d}$. It follows that

$$\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = \frac{\Psi_2}{\Psi_1} z I_t - \frac{1}{\Psi_1} \sum_{\boldsymbol{d}} \sum_{a=1}^{m} \sum_{b \neq a} \psi_a \psi_b y_{\boldsymbol{d},a,b} D_{\boldsymbol{d}},$$

as required. □

We continue by showing that $C_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})$ when the main effects of the attributes and contrasts of the position effects are of interest is block diagonal when using the designs introduced in Section 1.1.

■ **THEOREM 5.4.3.**

The $(1,1)$ *block of* $C(\boldsymbol{\pi}_0, \boldsymbol{\psi})$ *when using the generalised Davidson–Beaver position effects model, and the main effects of the attributes and contrasts of the position effects are of interest,* $C_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{MP}$*, is block diagonal.* □

Proof. Let P_{ℓ_q,e_q} be an $\ell_q \times \ell_q$ matrix with entries either 0 or 1 such that there is a 1 in position (t_1, t_2) if the difference between the two items is $t_2 - t_1 = e_q$. Then $P_{\ell_q,e_q} \otimes P_{\ell_q,e_q} \otimes \ldots \otimes P_{\ell_q,e_q}$ will give the pairs that have a difference $t_2 - t_1 = (e_1, e_2, \ldots, e_k)$. Let $\alpha_{\boldsymbol{e},a,b}$ be the number of times $\boldsymbol{e} = (e_1, e_2, \ldots, e_k)$ appears as a difference between the items in positions a and b of the choice set, and $\alpha_i = \sum_{e_1} \ldots \sum_{e_{i-1}} \sum_{e_{i+1}} \ldots \sum_{e_k} \alpha_{\boldsymbol{e}}$. Then

$$\begin{aligned}
\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) &= \frac{\Psi_2}{\Psi_1} z I_t - \frac{1}{\Psi_1} \sum_{\boldsymbol{d}} D_{\boldsymbol{d}} \sum_{a \neq b} \psi_a \psi_b y_{\boldsymbol{d},a,b} \\
&= \frac{1}{N\Psi_1} \Bigg[\Big(\sum_{e_1} \cdots \sum_{e_k} \sum_{a \neq b} \alpha_{\boldsymbol{e},a,b} \psi_a \psi_b \Big) \Big(P_{\ell_1,0} \otimes P_{\ell_2,0} \otimes \ldots \otimes P_{\ell_k,0} \Big) \\
&\quad - \sum_{e_1} \cdots \sum_{e_k} \sum_{a \neq b} \alpha_{\boldsymbol{e},a,b} \psi_a \psi_b \Big(P_{\ell_1,e_1} \otimes P_{\ell_2,e_2} \otimes \ldots \otimes P_{\ell_k,e_k} \Big) \Bigg].
\end{aligned}$$

Then the information matrix for the contrasts in $B_\gamma \boldsymbol{\gamma}$ is given by

$$\begin{aligned}
B_\gamma \Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) B_\gamma^T &= B_\gamma \frac{1}{N\Psi_1} \times \Bigg[\Big(\sum_{e_1} \cdots \sum_{e_k} \sum_{a \neq b} \alpha_{\boldsymbol{e},a,b} \psi_a \psi_b \Big) \Big(P_{\ell_1,0} \otimes P_{\ell_2,0} \otimes \ldots \otimes P_{\ell_k,0} \Big) \\
&\quad - \sum_{e_1} \cdots \sum_{e_k} \sum_{a \neq b} \alpha_{\boldsymbol{e},a,b} \psi_a \psi_b \Big(P_{\ell_1,e_1} \otimes P_{\ell_2,e_2} \otimes \ldots \otimes P_{\ell_k,e_k} \Big) \Bigg] B_\gamma^T \\
&= \frac{1}{N\Psi_1} \Bigg[\Big(\sum_{e_1} \cdots \sum_{e_k} \sum_{a \neq b} \alpha_{\boldsymbol{e},a,b} \psi_a \psi_b \Big) B_\gamma \Big(P_{\ell_1,0} \otimes P_{\ell_2,0} \otimes \ldots \otimes P_{\ell_k,0} \Big) B_\gamma^T \\
&\quad - \sum_{e_1} \cdots \sum_{e_k} \sum_{a \neq b} \alpha_{\boldsymbol{e},a,b} \psi_a \psi_b B_\gamma \Big(P_{\ell_1,e_1} \otimes P_{\ell_2,e_2} \otimes \ldots \otimes P_{\ell_k,e_k} \Big) B_\gamma^T \Bigg].
\end{aligned}$$

However, Burgess and Street [2005] and Street and Burgess [2007] showed that both

$$B_\gamma \Big(P_{\ell_1,0} \otimes P_{\ell_2,0} \otimes \ldots \otimes P_{\ell_k,0} \Big) B_\gamma^T \text{ and } B_\gamma \Big(P_{\ell_1,e_1} \otimes P_{\ell_2,e_2} \otimes \ldots \otimes P_{\ell_k,e_k} \Big) B_\gamma^T$$

are block diagonal matrices, so $C_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\text{MP}} = B_\gamma \Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) B_\gamma^T$ is also a block diagonal matrix. □

This theorem allows us to consider only the block diagonal entries of $C_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\text{MP}}$, that is, those which correspond to the main effects for a single attribute. In addition, Lemma 4.4.1 states that if each level of each attribute appears in each position of the choice set equally often then $C_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\text{MP}} = 0$, and therefore $C(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\text{MP}}$ is block diagonal.

The next theorem gives an expression for the block diagonal entry on $C_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\text{MP}}$ that corresponds to the main effects of attribute q.

■ **THEOREM 5.4.4.**

Under the null hypothesis of equal merits, the block diagonal entry of the information matrix corresponding to the main effect of attribute q is

$$\frac{1}{\Psi_1} \sum_{\boldsymbol{d}} \sum_{a \neq b} \psi_a \psi_b y_{\boldsymbol{d},a,b} \frac{\Gamma_{\boldsymbol{d}} \big((\ell_q - 1)^{i_q} - (-1)^{i_q} \big)}{(\ell_q - 1)^{i_q}} I_{\ell_q - 1},$$

where as before

$$\Gamma_{\boldsymbol{d}} = \prod_{q=1}^{k} (\ell_q - 1)^{i_q}.$$

□

Proof. By using

$$B_q D_{\boldsymbol{d}} B_q^T = \frac{\Gamma_{\boldsymbol{d}} (-1)^{i_q}}{(\ell_q - 1)^{i_q}} I_{\ell_q - 1},$$

as shown in Burgess and Street [2005], the q^{th} block of the block diagonal matrix $C_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\text{MP}}$ is given by

$$\begin{aligned}
B_q \Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) B_q^T &= B_q \left[\frac{\Psi_2}{\Psi_1} z I_t - \frac{1}{\Psi_1} \sum_{\boldsymbol{d}} D_{\boldsymbol{d}} \sum_{a \neq b} y_{\boldsymbol{d},a,b} \psi_a \psi_b \right] B_q^T \\
&= \frac{\Psi_2}{\Psi_1} z I_{\ell_q - 1} - \frac{1}{\Psi_1} \sum_{\boldsymbol{d}} \sum_{a \neq b} y_{\boldsymbol{d},a,b} \psi_a \psi_b \frac{\Gamma_{\boldsymbol{d}} (-1)^{i_q}}{(\ell_q - 1)^{i_q}} I_{\ell_q - 1},
\end{aligned}$$

where B_q is the matrix containing the contrast coefficients for the main effects of the q^{th} attribute.

By substituting in the expression for z, $C_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\text{MP}}$ can be expressed as

$$\begin{aligned} B_q \Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) B_q^T &= \frac{\Psi_2}{\Psi_1} \sum_{\boldsymbol{d}} \frac{1}{N} \sum_j c_{\boldsymbol{v}_j} i_{\boldsymbol{v}_j} x_{\boldsymbol{v}_j;\boldsymbol{d},a,b} I_{\ell_q - 1} - \frac{1}{\Psi_1} \sum_{\boldsymbol{d}} \sum_{a \neq b} y_{\boldsymbol{d},a,b} \psi_a \psi_b \frac{\Gamma_{\boldsymbol{d}} (-1)^{i_q}}{(\ell_q - 1)^{i_q}} I_{\ell_q - 1} \\ &= \frac{1}{\Psi_1} \sum_{\boldsymbol{d}} \Gamma_{\boldsymbol{d}} \sum_{a \neq b} \psi_a \psi_b y_{\boldsymbol{d},a,b} I_{\ell_q - 1} - \frac{1}{\Psi_1} \sum_{\boldsymbol{d}} \sum_{a \neq b} y_{\boldsymbol{d},a,b} \psi_a \psi_b \frac{\Gamma_{\boldsymbol{d}} (-1)^{i_q}}{(\ell_q - 1)^{i_q}} I_{\ell_q - 1} \\ &= \frac{1}{\Psi_1} \sum_{\boldsymbol{d}} \sum_{a \neq b} \psi_a \psi_b y_{\boldsymbol{d},a,b} \frac{\Gamma_{\boldsymbol{d}} \big((\ell_q - 1)^{i_q} - (-1)^{i_q} \big)}{(\ell_q - 1)^{i_q}} I_{\ell_q - 1}, \end{aligned}$$

as required. □

We can use this theorem to give an expression for the determinant of the information matrix when the generalised Davidson–Beaver position effects model is used and both the main effects of the attributes and contrasts of the position effects are of interest.

■ **THEOREM 5.4.5.**

When each level of each attribute appears equally often across each of the positions, the determinant of the information matrix for the estimation of the main effects of the attributes and contrasts of the position effects under the null hypothesis of equal merits, when the Davidson–Beaver position effects model is used, is given by

$$\det(C(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{MP}) = \prod_{q=1}^{k} \left[\frac{\ell_q}{N \Psi_1 (\ell_q - 1)} \sum_j c_{\boldsymbol{v}_j} i_{\boldsymbol{v}_j} \sum_{a \neq b} \psi_a \psi_b \sum_{\boldsymbol{d} | i_q = 1} x_{\boldsymbol{v}_j;\boldsymbol{d},a,b} \right]^{\ell_q - 1} \times \det(C_{\psi\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}))$$

where $\det(C_{\psi\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}))$ *is the determinant of the* $(2,2)$ *block of the information matrix, and is independent of the design chosen for a given choice set size.* □

Proof. Since the $(1,1)$ block of the information matrix for the estimation of the main effects of the attributes and contrasts of the position effects is block diagonal, as shown in Theorem 5.4.4, the determinant of the information matrix will be the product of the determinants of each block of the information matrix. This gives

$$\begin{aligned} \det(C_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})) &= \prod_{q=1}^{k} \left[\frac{1}{\Psi_1} \sum_{\boldsymbol{d}} \sum_{a \neq b} \psi_a \psi_b y_{\boldsymbol{d},a,b} \frac{\Gamma_{\boldsymbol{d}} \big((\ell_q - 1)^{i_q} - (-1)^{i_q} \big)}{(\ell_q - 1)^{i_q}} \right]^{\ell_q - 1} \\ &= \prod_{q=1}^{k} \left[\frac{1}{N\Psi_1} \sum_{\boldsymbol{d}} \sum_{a \neq b} \psi_a \psi_b \frac{1}{\Gamma_{\boldsymbol{d}}} \sum_j c_{\boldsymbol{v}_j} i_{\boldsymbol{v}_j} x_{\boldsymbol{v}_j;\boldsymbol{d},a,b} \frac{\Gamma_{\boldsymbol{d}} \big((\ell_q - 1)^{i_q} - (-1)^{i_q} \big)}{(\ell_q - 1)^{i_q}} \right]^{\ell_q - 1} \\ &= \prod_{q=1}^{k} \left[\frac{1}{N\Psi_1} \sum_{\boldsymbol{d}} \sum_{a \neq b} \sum_j \frac{\psi_a \psi_b}{(\ell_q - 1)^{i_q}} c_{\boldsymbol{v}_j} i_{\boldsymbol{v}_j} x_{\boldsymbol{v}_j;\boldsymbol{d},a,b} \big((\ell_q - 1)^{i_q} - (-1)^{i_q} \big) \right]^{\ell_q - 1} \\ &= \prod_{q=1}^{k} \left[\frac{\ell_q}{N \Psi_1 (\ell_q - 1)} \sum_j c_{\boldsymbol{v}_j} i_{\boldsymbol{v}_j} \sum_{a \neq b} \psi_a \psi_b \sum_{\boldsymbol{d} | i_q = 1} x_{\boldsymbol{v}_j;\boldsymbol{d},a,b} \right]^{\ell_q - 1}. \end{aligned}$$

So when each level in each attribute appears in each position equally often, the determinant of the information matrix for the estimation of the main effects of the attributes and contrasts of the position effects is

$$\det(C(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\text{MP}}) = \prod_{q=1}^{k} \left[\frac{\ell_q}{N \Psi_1 (\ell_q - 1)} \sum_j c_{\boldsymbol{v}_j} i_{\boldsymbol{v}_j} \sum_{a \neq b} \psi_a \psi_b \sum_{\boldsymbol{d} | i_q = 1} x_{\boldsymbol{v}_j;\boldsymbol{d},a,b} \right]^{\ell_q - 1} \times \det(C_{\psi\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi})).$$

Since $\Lambda_{\psi\psi}(\boldsymbol{\pi}_0,\boldsymbol{\psi})$ depends only on the value of m and the entries in $\boldsymbol{\psi}$, $C_{\psi\psi}(\boldsymbol{\pi}_0,\boldsymbol{\psi})$ is also a function of m and the entries in $\boldsymbol{\psi}$ only. That is, $C_{\psi\psi}(\boldsymbol{\pi}_0,\boldsymbol{\psi})$ is independent of the design chosen, for a fixed value of m under the null hypothesis of equal merits. □

We can use this expression to extend the result in Burgess and Street [2005], to find the optimum value of the determinant of the information matrix for the estimation of the main effects of the attributes and contrasts of the position effects when the generalised Davidson–Beaver position effects model is used.

■ **THEOREM 5.4.6.**

The D–optimal design for the estimation of the main effects of the attributes and contrasts of the position effects will be given by the set of choice sets where at least one difference vector $\boldsymbol{v}_j$ has a non–zero $a_{\boldsymbol{v}_j}$, and for each $\boldsymbol{v}_j$ present, and for each attribute q,

$$S_q = \begin{cases} \frac{m^2-1}{4}, & \ell_q = 2 \text{ and } m \text{ is odd;} \\ \frac{m^2}{4}, & \ell_q = 2 \text{ and } m \text{ is even;} \\ \frac{m^2-(\ell_q x^2+2xy+y)}{2}, & 2 < \ell_q < m; \\ \frac{m(m-1)}{2}, & \ell_q \geq m, \end{cases}$$

where positive integers x and y satisfy the equation $m = \ell_q x + y$ for $0 \leq y < \ell_q - 1$. The maximum possible value for the determinant of the information matrix will be

$$\det((C(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{MP})_{OPT}) = \prod_{q=1}^{k}\left[\frac{2S_q\ell_q\Psi_2}{Lm(m-1)\Psi_1(\ell_q-1)}\right]^{\ell_q-1} \times \det(C_{\psi\psi}(\boldsymbol{\pi}_0,\boldsymbol{\psi})),$$

given that each pair of positions contains equally many non–zero differences. □

Proof. In order to maximise the information matrix, $\sum_{a\neq b}\psi_a\psi_b\sum_{\boldsymbol{d}|i_q=1}x_{\boldsymbol{v}_j;\boldsymbol{d},a,b}$ will need to be maximised. Under the condition that each pair of positions contains equally many non–zero differences, we obtain

$$\sum_{a\neq b}\psi_a\psi_b\sum_{\boldsymbol{d}|i_q=1}x_{\boldsymbol{v}_j;\boldsymbol{d},a,b} = \frac{2}{m(m-1)}\sum_{\boldsymbol{d}|i_q=1}x_{\boldsymbol{v}_j;\boldsymbol{d}} \times \Psi_2,$$

where Ψ_2 is independent of the design used. Theorem 1 in Burgess and Street [2005] shows that $\sum_{\boldsymbol{d}|i_q=1}x_{\boldsymbol{v}_j;\boldsymbol{d}}$ is maximised when it is equal to S_q, where S_q is defined in the statement of the theorem. In this case, the determinant of the $(1,1)$ block of the information matrix, $C_{\gamma\gamma}(\boldsymbol{\pi}_0,\boldsymbol{\psi})$ relating to the information provided about the main effects of the attributes, will be

$$\begin{aligned} \det(C_{\gamma\gamma}(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\text{OPT}}) &= \prod_{q=1}^{k}\left[\frac{2S_q\ell_q\Psi_2}{m(m-1)\Psi_1 L\sum_j c_{\boldsymbol{v}_j}i_{\boldsymbol{v}_j}(\ell_q-1)}\sum_j c_{\boldsymbol{v}_j}i_{\boldsymbol{v}_j}\right]^{\ell_q-1} \\ &= \prod_{q=1}^{k}\left[\frac{2S_q\ell_q\Psi_2}{m(m-1)\Psi_1 L(\ell_q-1)}\right]^{\ell_q-1}. \end{aligned}$$

Since all of the ordered choice sets with a particular difference vector are assumed to be included equally often, each attribute level will appear in each position equally often, and therefore $C_{\gamma\psi}(\boldsymbol{\pi}_0,\boldsymbol{\psi}) = \mathbf{0}$. For a given m, $C_{\psi\psi}(\boldsymbol{\pi}_0,\boldsymbol{\psi})$ is fixed across all designs, and then for all $\ell_1,\ell_2,\ldots\ell_k$, and k, the optimal information matrix for the estimation of the main effects of the attributes will be

$$\det((C(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\mathrm{MP}})_{\mathrm{OPT}}) \quad = \quad \prod_{q=1}^{k}\left[\frac{2S_q\ell_q\Psi_2}{Lm(m-1)\Psi_1(\ell_q-1)}\right]^{\ell_q-1} \times \det(C_{\psi\psi}(\boldsymbol{\pi}_0,\boldsymbol{\psi})),$$

as required. □

We now look at an example where we compare the determinant of the information matrix for the estimation of the main effects of the attributes plus contrasts of the position effects to the optimal determinant given in the previous theorem.

■ **EXAMPLE 5.4.2.**

Consider an experiment with two 3-level attributes presented in choice sets of size 3. Suppose that the choice experiment consists of all choice sets with the difference vector $(11, 11, 11)$. This gives the design in Table 5.5. The first two items in each choice set differ in both attributes, as do the first and third items in each choice set, and the second and third items in each choice set. Then each pair of positions have 2 non–zero differences.

Option1	**Option2**	**Option3**	**Option1**	**Option2**	**Option3**
0 0	1 1	2 2	0 0	2 2	1 1
0 1	1 2	2 0	0 1	2 0	1 2
0 2	1 0	2 1	0 2	2 1	1 0
1 0	2 1	0 2	1 0	0 2	2 1
1 1	2 2	0 0	1 1	0 0	2 2
1 2	2 0	0 1	1 2	0 1	2 0
2 0	0 1	1 2	2 0	1 2	0 1
2 1	0 2	1 0	2 1	1 0	0 2
2 2	0 0	1 1	2 2	1 1	0 0
0 0	1 2	2 1	0 0	2 1	1 2
0 1	1 0	2 2	0 1	2 2	1 0
0 2	1 1	2 0	0 2	2 0	1 1
1 0	2 2	0 1	1 0	0 1	2 2
1 1	2 0	0 2	1 1	0 2	2 0
1 2	2 1	0 0	1 2	0 0	2 1
2 0	0 2	1 1	2 0	1 1	0 2
2 1	0 0	1 2	2 1	1 2	0 0
2 2	0 1	1 0	2 2	1 0	0 1

Table 5.5: The design for the choice experiment in Example 5.4.2

The information matrix for the estimation of the main effects of the attributes plus contrasts of the position effects is shown in Figure 5.1, and has determinant

$$\det(C(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\mathrm{MP}}) = \frac{(\psi_1\psi_2+\psi_1\psi_3+\psi_2\psi_3)^{4}}{243\psi_1\psi_2\psi_3\,(\psi_1+\psi_2+\psi_3)^{9}}.$$

$$C(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\mathrm{MP}} =$$

$$\begin{bmatrix}
\frac{\psi_2\psi_3+\psi_1(\psi_2+\psi_3)}{3(\psi_1+\psi_2+\psi_3)^2} & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{\psi_2\psi_3+\psi_1(\psi_2+\psi_3)}{3(\psi_1+\psi_2+\psi_3)^2} & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{\psi_2\psi_3+\psi_1(\psi_2+\psi_3)}{3(\psi_1+\psi_2+\psi_3)^2} & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{\psi_2\psi_3+\psi_1(\psi_2+\psi_3)}{3(\psi_1+\psi_2+\psi_3)^2} & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{\psi_1+\psi_3}{2\psi_1\psi_3(\psi_1+\psi_2+\psi_3)} & \frac{\psi_1-\psi_3}{2\sqrt{3}\psi_1\psi_3(\psi_1+\psi_2+\psi_3)} \\
0 & 0 & 0 & 0 & \frac{\psi_1-\psi_3}{2\sqrt{3}\psi_1\psi_3(\psi_1+\psi_2+\psi_3)} & \frac{\psi_2\psi_3+\psi_1(\psi_2+4\psi_3)}{6\psi_1\psi_2\psi_3(\psi_1+\psi_2+\psi_3)}
\end{bmatrix}$$

Figure 5.1: The information matrix for the design in Example 5.4.2

We now compare the determinant above to the optimal determinant as given by Theorem 5.4.6. Since $\ell_1 = \ell_2 = m = 3$, we have

$$S_1 = S_2 = \frac{m(m-1)}{2} = 3,$$

and

$$\det(C_{\psi\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi})) = \frac{1}{3\psi_1\psi_2\psi_3(\psi_1\psi_2\psi_3)}.$$

Then

$$\begin{aligned}\prod_{q=1}^{k}\left(\frac{2S_q\ell_q\Psi_2}{Lm(m-1)\Psi_1(\ell_q-1)}\right)^{\ell_q-1} &= \prod_{q=1}^{2}\left(\frac{4\times 3\times 3\Psi_2}{9\times 3\times 2\times \Psi_1\times 2}\right)^2 \\ &= \left(\frac{\Psi_2}{6\Psi_1}\right)^2 \times \left(\frac{\Psi_2}{6\Psi_1}\right)^2 \\ &= \frac{\Psi_2^4}{1296\Psi_1^4}.\end{aligned}$$

Then the optimum value of the determinant of the information matrix for the estimation of the main effects of the attributes and contrasts of the position effects is

$$\begin{aligned}\det((C(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\mathrm{MP}})_{\mathrm{OPT}}) &= \frac{\Psi_2^4}{1296\Psi_1^4} \times \frac{1}{3\psi_1\psi_2\psi_3(\psi_1\psi_2\psi_3)} \\ &= \frac{(\psi_1\psi_2+\psi_1\psi_3+\psi_2\psi_3)^4}{243\psi_1\psi_2\psi_3\,(\psi_1+\psi_2+\psi_3)^9}.\end{aligned}$$

Since this is equal to the optimal value of the determinant of the information matrix for the estimation of the main effects of the attributes and contrasts of the position effects, this design is optimal for the estimation of the main effects of the attributes and contrasts of the position effects when using the Davidson–Beaver position effects model. □

The expression in Theorem 5.4.6 allows us to confirm the optimal designs for the estimation of the main effects of the attributes and contrasts of the position effects when using the generalised Davidson–Beaver position effects model.

■ **THEOREM 5.4.7.**

Let F be the complete factorial for k attributes where the q^{th} attribute has ℓ_q levels. Suppose that we choose a set of m generators $G = \{\boldsymbol{g}_1 = \mathbf{0}, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_m\}$ such that $\boldsymbol{g}_i \neq \boldsymbol{g}_j$ for $i \neq j$. Suppose that $\boldsymbol{g}_i = (g_{i1}, g_{i2}, \ldots, g_{ik})$ for $i = 1, \ldots, m$ and also that the multiset of differences for attribute q,

$$\{\pm(g_{i_1q} - g_{i_2q})|1 \leq i_1, i_2 \leq m, i_1 \neq i_2\},$$

contains each non-zero difference modulo ℓ_q equally often. Then the ordered choice sets given by the rows of $F+\boldsymbol{g}_1, F+\boldsymbol{g}_2, \ldots, F+\boldsymbol{g}_m$, for one or more sets of generators G, are optimal for the estimation of the main effects of the attributes and contrasts of the position effects, provided that there are as few zero differences as possible in each choice set, each pair of positions contains equally many non–zero differences, and that each level of an attribute is equally replicated across each position of the set of ordered choice sets. □

Proof. Theorem 5.4.4 showed that the $(1,1)$ block of $C(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\mathrm{MP}}$ can be written as

$$B_q\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})B_q^T \quad = \quad \frac{1}{\Psi_1}\sum_{\boldsymbol{d}}\sum_{a\neq b}\psi_a\psi_b y_{\boldsymbol{d},a,b}\frac{\Gamma_{\boldsymbol{d}}\left((\ell_q-1)^{i_q}-(-1)^{i_q}\right)}{(\ell_q-1)^{i_q}}I_{\ell_q-1}.$$

Substituting the expression for $y_{\boldsymbol{d},a,b}$ given earlier, and simplifying, gives

$$\begin{aligned} B_q\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0,\boldsymbol{\psi})B_q^T &= \frac{\ell_q}{N\Psi_1(\ell_q-1)}\sum_{\boldsymbol{d}|i_q=1}\sum_{a\neq b}\psi_a\psi_b\sum_j c_{\boldsymbol{v}_j}i_{\boldsymbol{v}_j}x_{\boldsymbol{v}_j;\boldsymbol{d},a,b}\Gamma_{\boldsymbol{d}}I_{\ell_q-1} \\ &= \frac{\ell_q}{N\Psi_1(\ell_q-1)}\sum_j c_{\boldsymbol{v}_j}i_{\boldsymbol{v}_j}\sum_{a\neq b}\sum_{\boldsymbol{d}|i_q=1}\psi_a\psi_b x_{\boldsymbol{v}_j;\boldsymbol{d},a,b}I_{\ell_q-1}. \end{aligned}$$

Using the assumption that the differences between any two positions contains all of the non–zero differences equally often, we obtain

$$\sum_{a\neq b}x_{\boldsymbol{v}_j;\boldsymbol{d},a,b}\psi_a\psi_b = \frac{2}{m(m-1)}\alpha_q\Psi_2,$$

where α_q is the number of non–zero differences occurring in the q^{th} position of the differences in the difference vectors. Substituting this into $B_q\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0,\boldsymbol{\psi})B_q^T$ gives

$$B_q\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0,\boldsymbol{\psi})B_q^T = \frac{2\ell_q\alpha_q\Psi_2}{N\Psi_1 m(m-1)(\ell_q-1)}\sum_j c_{\boldsymbol{v}_j}i_{\boldsymbol{v}_j}I_{\ell_q-1}.$$

However, $\sum_j c_{\boldsymbol{v}_j}i_{\boldsymbol{v}_j}=\ell_q-1$, hence

$$B_q\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0,\boldsymbol{\psi})B_q^T = \frac{2\ell_q\alpha_q\Psi_2}{N\Psi_1 m(m-1)}I_{\ell_q-1}.$$

Now that we have shown that the q^{th} diagonal block for the matrix $C_{\gamma\gamma}(\boldsymbol{\pi}_0,\boldsymbol{\psi})$ is equal to

$$\frac{2\ell_q\alpha_q\Psi_2}{N\Psi_1 m(m-1)}I_{\ell_q-1},$$

we can see that the determinant of this block matrix will be maximised if α_q is maximised.

Since we have the condition that each non–zero difference must occur equally often, there are $(\ell_q-1)\alpha_q$ generators with non–zero differences in attribute q. Also, since the complete factorial has been used as the initial block, there will be $L(\ell_q-1)\alpha_q$ non–zero differences for attribute q in the experiment. So,

$$\begin{aligned} L(\ell_q-1)\alpha_q &= S_qN \\ \frac{(\ell_q-1)\alpha_q}{N} &= \frac{S_q}{L} \\ \frac{2\ell_q\alpha_q\Psi_2}{N\Psi_1 m(m-1)} &= \frac{2\ell_q S_q\Psi_2}{L\Psi_1 m(m-1)(\ell_q-1)}. \end{aligned}$$

This is the same as the entry for attribute q in the expression in Theorem 5.4.6, and therefore this design is optimal for the estimation of the main effects of the attributes and contrasts of the position effects. □

Example 5.4.2 illustrates these results.

5.4.2 Main Effects plus Two–Factor Interactions and Position Effects

In this subsection, we extend the results developed for the estimation of the main effects of the attributes and contrasts of the position effects when the generalised Davidson–Beaver position effects model is used to the case where we are interested in the estimation of two–factor interactions between the attributes in addition to the main effects of the attributes and contrasts of the position effects. In this subsection, we will consider criteria for independence of attribute main effects and two–factor interactions from position effects. We then prove a theorem that gives an expression for the information matrix when attribute main effects plus two–factor interactions and contrasts of the position effects are all of interest and the generalised Davidson–Beaver position effects model is used.

We now prove a lemma that gives conditions for $C_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = \mathbf{0}$, when attribute main effects plus two–factor interactions and contrasts of the position effects are of interest.

■ **LEMMA 5.4.8.**

The information matrix for the estimation of the main effects plus two–factor interactions of the attributes and contrasts of the position effects is block diagonal with respect to the attribute effects and the position effect if, for each pair of attributes, each pair of attribute levels appears equally often in each position of the choice set.

Proof. Suppose that B_{MT} contains the contrast coefficients corresponding to the main effects plus two–factor interactions of the attributes. The contrast coefficients corresponding to items with the same levels in the pair of attributes that are of interest will be the same. Then for the j^{th} contrast, corresponding to a component of the two–factor interaction between attributes q_1 and q_2,

$$(B_{\mathrm{MT}}\Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}))_{ja} = \frac{1}{\Psi_1}\sum_{b\neq a}\psi_b \sum_{x_1=1}^{\ell_{q_1}}\sum_{x_2=1}^{\ell_{q_2}} B_{j(x_1x_2)}(\lambda_{q_1=x_1,q_2=x_2 \text{ in pos } a} - \lambda_{q_1=x_1,q_2=x_2 \text{ in pos } b}),$$

where $\lambda_{q_1=x_1,q_2=x_2 \text{ in pos } a} = \lambda_{\text{att } q_1=x_1, \text{ att } q_2=x_2 \text{ in pos } a}$. Then $(B_{\mathrm{MT}}\Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}))_{ja} = 0$ if, for all x_1 and x_2,

$$\lambda_{q_1=x_1,q_2=x_2 \text{ in pos } a} - \lambda_{q_1=x_1,q_2=x_2 \text{ in pos } b} = 0,$$

for all $b \neq a$. If this is the case, then it follows that $(B_{\mathrm{MT}}\Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}))_{ja} = 0$, where contrast j now corresponds to a main effect of either attribute q_1 or attribute q_2, as each level of the attribute must appear in each position of the choice set equally often if each pair of attribute levels appears in each position equally often. It follows that $B_{\mathrm{MT}}\Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = \mathbf{0}$ if, for each pair of attributes, each pair of attribute levels appears equally often in each of the positions in the choice set. Then the information matrix for the estimation of the main effects plus two–factor interactions of the attributes and contrasts of the position effects is block diagonal, with one block corresponding to the attribute effects and one block corresponding to the position effects. □

■ **EXAMPLE 5.4.3.**

Consider the experiment discussed in Example 5.4.2, and the design in Table 5.5. In this choice experiment, each item appears in each position of the choice set four times. Thus this design satisfies the criteria for Lemma 5.4.8. $B_{\mathrm{MT}}\Lambda(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = \mathbf{0}$ when this design is used as can be seen in Figure 5.2. Notice that this matrix is block diagonal. □

We can use this lemma to prove a theorem that gives the determinant of the information matrix for the estimation of the main effects plus two–factor interactions of the attributes and contrasts of the position effects when the generalised Davidson–Beaver position effects model is used.

■ **THEOREM 5.4.9.**

Under the null hypothesis of equal merits, the determinant of the information matrix for the

$$B_{\mathrm{MT}}\Lambda(\boldsymbol{\pi}_0, \boldsymbol{\psi}) =$$

$$\begin{bmatrix}
-\frac{\Psi_2}{6\sqrt{6}\Psi_1} & -\frac{\Psi_2}{6\sqrt{6}\Psi_1} & -\frac{\Psi_2}{6\sqrt{6}\Psi_1} & 0 & 0 & 0 & \frac{\Psi_2}{6\sqrt{6}\Psi_1} & \frac{\Psi_2}{6\sqrt{6}\Psi_1} & \frac{\Psi_2}{6\sqrt{6}\Psi_1} & 0 & 0 & 0 \\
\frac{\Psi_2}{18\sqrt{2}\Psi_1} & \frac{\Psi_2}{18\sqrt{2}\Psi_1} & \frac{\Psi_2}{18\sqrt{2}\Psi_1} & -\frac{\sqrt{2}\Psi_2}{18\Psi_1} & -\frac{\sqrt{2}\Psi_2}{18\Psi_1} & -\frac{\sqrt{2}\Psi_2}{18\Psi_1} & \frac{\Psi_2}{18\sqrt{2}\Psi_1} & \frac{\Psi_2}{18\sqrt{2}\Psi_1} & \frac{\Psi_2}{18\sqrt{2}\Psi_1} & 0 & 0 & 0 \\
-\frac{\Psi_2}{6\sqrt{6}\Psi_1} & 0 & \frac{\Psi_2}{6\sqrt{6}\Psi_1} & -\frac{\Psi_2}{6\sqrt{6}\Psi_1} & 0 & \frac{\Psi_2}{6\sqrt{6}\Psi_1} & -\frac{\Psi_2}{6\sqrt{6}\Psi_1} & 0 & \frac{\Psi_2}{6\sqrt{6}\Psi_1} & 0 & 0 & 0 \\
\frac{\Psi_2}{9\sqrt{2}\Psi_1} & -\frac{\sqrt{2}\Psi_2}{18\Psi_1} & \frac{\Psi_2}{18\sqrt{2}\Psi_1} & \frac{\Psi_2}{18\sqrt{2}\Psi_1} & -\frac{\sqrt{2}\Psi_2}{18\Psi_1} & \frac{\Psi_2}{18\sqrt{2}\Psi_1} & \frac{\Psi_2}{18\sqrt{2}\Psi_1} & -\frac{\sqrt{2}\Psi_2}{18\Psi_1} & \frac{\Psi_2}{18\sqrt{2}\Psi_1} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{2}\psi_1\sqrt{\Psi_1}} & 0 & \frac{1}{\sqrt{2}\psi_3\sqrt{\Psi_1}} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{\sqrt{6}\psi_1\sqrt{\Psi_1}} & -\frac{\sqrt{\frac{2}{3}}}{\psi_2\sqrt{\Psi_1}} & \frac{1}{\sqrt{6}\psi_3\sqrt{\Psi_1}}
\end{bmatrix}$$

Figure 5.2: $B_{\mathrm{MT}}\Lambda(\boldsymbol{\pi}_0, \boldsymbol{\psi})$ for the design in Table 5.5

estimation of the main effects plus two–factor interactions of the attributes, and contrasts of the position effects, when the generalised Davidson–Beaver position effects model is used, is given by

$$\begin{aligned}
&\det(C(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{MTP}) \\
&= \prod_{q=1}^{k}\left[\frac{1}{\Psi_1}\frac{\ell_q}{(\ell_q-1)}\sum_j c_{\boldsymbol{v}_j}a_{\boldsymbol{v}_j}\sum_{a\neq b}\psi_a\psi_b\sum_{\boldsymbol{d}|i_q=1}x_{\boldsymbol{v}_j;\boldsymbol{d},a,b}\right]^{\ell_q-1} \\
&\times\prod_{q_1=1}^{k-1}\prod_{q_2=q_1+1}^{k}\left[\frac{\Psi_2}{\Psi_1}\sum_{\boldsymbol{d}}\sum_{a\neq b}\sum_j\psi_a\psi_b c_{\boldsymbol{v}_j}a_{\boldsymbol{v}_j}x_{\boldsymbol{v}_j;\boldsymbol{d},a,b}\right. \\
&\times\left.\left[1-\frac{1}{(1-\ell_{q_1})^{i_{q_1}}(1-\ell_{q_2})^{i_{q_2}}}\right]I_{(\ell_{q_1}-1)(\ell_{q_2}-1)}\right]^{(\ell_{q_1}-1)(\ell_{q_2}-1)} \times\det(C_{\psi\psi}(\boldsymbol{\pi},\boldsymbol{\psi})),
\end{aligned}$$

when each pair of attribute levels appears equally often in each position in the set of ordered choice sets. □

Proof. The first step in the proof is to show that the $(1,1)$ block of the information matrix $C(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\text{MTP}}$ is block diagonal. We begin with

$$\begin{aligned}
C_{\gamma\gamma}(\boldsymbol{\pi}_0,\boldsymbol{\psi}) &= B_{MT}\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0,\boldsymbol{\psi})B_{MT}^T \\
&= B_{MT}\left[\frac{\Psi_2}{\Psi_1}zI_L-\frac{1}{\Psi_1}\sum_{\boldsymbol{d}}D_{\boldsymbol{d}}\sum_{a\neq b}\psi_a\psi_b y_{\boldsymbol{d},a,b}\right]B_{MT}^T \\
&= \frac{\Psi_2}{\Psi_1}zB_{MT}I_LB_{MT}^T-\frac{1}{\Psi_1}\sum_{\boldsymbol{d}}\sum_{a\neq b}\psi_a\psi_b y_{\boldsymbol{d},a,b}B_{MT}D_{\boldsymbol{d}}B_{MT}^T,
\end{aligned}$$

where B_{MT} contains the contrast coefficients for the main effects and the two factor interactions of the attributes. Burgess and Street [2005] have shown that

$$B_{MT}D_{\boldsymbol{d}}B_{MT}^T=\begin{bmatrix}B_MD_{\boldsymbol{d}}B_M^T & 0\\ 0 & B_TD_{\boldsymbol{d}}B_T^T\end{bmatrix},$$

where the $(\ell_q-1)\times(\ell_q-1)$ block matrix of $B_MD_{\boldsymbol{d}}B_M^T$ for attribute q is

$$\frac{\Gamma_{\boldsymbol{d}}(-1)^{i_q}}{(\ell_q-1)^{i_q}}I_{\ell_q-1},$$

and for the two–factor interactions, the $(\ell_{q_1}-1)(\ell_{q_2}-1)\times(\ell_{q_1}-1)(\ell_{q_2}-1)$ block matrix of $B_TD_{\boldsymbol{d}}B_T^T$ for attributes q_1 and q_2 is

$$\frac{\Gamma_{\boldsymbol{d}}(-1)^{i_{q_1}}(-1)^{i_{q_2}}}{(\ell_{q_1}-1)^{i_{q_1}}(\ell_{q_2}-1)^{i_{q_2}}}I_{(\ell_{q_1}-1)(\ell_{q_2}-1)}. \tag{5.6}$$

By using this result, and the orthogonality property of the contrast matrix, we obtain

$$C_{\gamma\gamma}(\boldsymbol{\pi}_0,\boldsymbol{\psi}) = \frac{\Psi_2}{\Psi_1}zI_p-\frac{1}{\Psi_1}\sum_{\boldsymbol{d}}\sum_{a\neq b}\psi_a\psi_b y_{\boldsymbol{d},a,b}\begin{bmatrix}B_MD_{\boldsymbol{d}}B_M^T & 0\\ 0 & B_TD_{\boldsymbol{d}}B_T^T\end{bmatrix}.$$

Hence

$$\begin{aligned}
C_{\gamma\gamma}(\boldsymbol{\pi}_0,\boldsymbol{\psi}) = \text{BlkDiag}\Bigg(&\frac{\Psi_2}{\Psi_1}zI_{\prod_{q=1}^k(\ell_q-1)}-\frac{1}{\Psi_1}\sum_{\boldsymbol{d}}\sum_{a\neq b}\psi_a\psi_b y_{\boldsymbol{d},a,b}B_MD_{\boldsymbol{d}}B_M^T, \\
&\frac{\Psi_2}{\Psi_1}zI_{\prod_{q_1\neq q_2}(\ell_{q_1}-1)(\ell_{q_2}-1)}-\frac{1}{\Psi_1}\sum_{\boldsymbol{d}}\sum_{a\neq b}\psi_a\psi_b y_{\boldsymbol{d},a,b}B_TD_{\boldsymbol{d}}B_T^T\Bigg).
\end{aligned}$$

So clearly the $(1,1)$ block of $C(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\mathrm{MTP}}$ is block diagonal.

Next, we find an expression for the $(\ell_{q_1}-1)(\ell_{q_2}-1)\times(\ell_{q_1}-1)(\ell_{q_2}-1)$ block matrix corresponding to the two–factor interactions between attributes q_1 and q_2. Let $B_{T_{q_1,q_2}}$ contain the contrast coefficients for the two–factor interactions involving attributes q_1 and q_2. Then we substitute Equation 5.6, and simplify, to obtain

$$\begin{aligned}
&\frac{\Psi_2}{\Psi_1} z I_{(\ell_{q_1}-1)(\ell_{q_2}-1)} - \frac{1}{\Psi_1}\sum_{\boldsymbol{d}}\sum_{a\neq b}\psi_a\psi_b y_{\boldsymbol{d},a,b} B_{T_{q_1,q_2}} D_{\boldsymbol{d}} B_{T_{q_1,q_2}}^T \\
&= \frac{\Psi_2}{\Psi_1} z I_{(\ell_{q_1}-1)(\ell_{q_2}-1)} - \frac{1}{\Psi_1}\sum_{\boldsymbol{d}}\sum_{a\neq b}\psi_a\psi_b y_{\boldsymbol{d},a,b}\frac{\Gamma_{\boldsymbol{d}}(-1)^{i_{q_1}+i_{q_2}}}{(\ell_{q_1}-1)^{i_{q_1}}(\ell_{q_2}-1)^{i_{q_2}}} I_{(\ell_{q_1}-1)(\ell_{q_2}-1)} \\
&= \frac{1}{N\Psi_1}\sum_{\boldsymbol{d}}\sum_{a\neq b}\Gamma_{\boldsymbol{d}}\psi_a\psi_b\frac{1}{\Gamma_{\boldsymbol{d}}}\times\sum_j c_{\boldsymbol{v}_j} i_{\boldsymbol{v}_j} x_{\boldsymbol{v}_j;\boldsymbol{d},a,b}\left[1-\frac{1}{(1-\ell_{q_1})^{i_{q_1}}(1-\ell_{q_2})^{i_{q_2}}}\right] I_{(\ell_{q_1}-1)(\ell_{q_2}-1)} \\
&= \frac{1}{N\Psi_1}\sum_{\boldsymbol{d}}\sum_{a\neq b}\sum_j \psi_a\psi_b c_{\boldsymbol{v}_j} i_{\boldsymbol{v}_j} x_{\boldsymbol{v}_j;\boldsymbol{d},a,b}\left[1-\frac{1}{(1-\ell_{q_1})^{i_{q_1}}(1-\ell_{q_2})^{i_{q_2}}}\right] I_{(\ell_{q_1}-1)(\ell_{q_2}-1)}.
\end{aligned}$$

Lemma 5.4.8 showed that the main effects and two–factor interactions of the attributes will be independent of position effects. Using this, and the result in Theorem 5.4.5, the determinant for the information matrix of the estimation for main effects, two–factor interactions, and position effects is

$$\begin{aligned}
\det(C(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\mathrm{MTP}}) &= \prod_{q=1}^{k}\left[\frac{\ell_q}{N\Psi_1(\ell_q-1)}\sum_j c_{\boldsymbol{v}_j} i_{\boldsymbol{v}_j}\sum_{a\neq b}\psi_a\psi_b\sum_{\boldsymbol{d}|i_q=1} x_{\boldsymbol{v}_j;\boldsymbol{d},a,b}\right]^{\ell_q-1} \\
&\quad\times\prod_{q_1=1}^{k-1}\prod_{q_2=q_1+1}^{k}\left[\frac{1}{N\Psi_1}\sum_{\boldsymbol{d}}\sum_{a\neq b}\sum_j \psi_a\psi_b c_{\boldsymbol{v}_j} i_{\boldsymbol{v}_j} x_{\boldsymbol{v}_j;\boldsymbol{d},a,b}\right. \\
&\qquad\left.\left[1-\frac{1}{(1-\ell_{q_1})^{i_{q_1}}(1-\ell_{q_2})^{i_{q_2}}}\right] I_{(\ell_{q_1}-1)(\ell_{q_2}-1)}\right]^{(\ell_{q_1}-1)(\ell_{q_2}-1)} \\
&\quad\times\det(C_{\psi\psi}(\boldsymbol{\pi}_0,\boldsymbol{\psi})),
\end{aligned}$$

as required. □

We now demonstrate this theorem with an example.

■ **EXAMPLE 5.4.4.**

Consider again the experiment in Example 5.4.2 and the design in Table 5.5. A generic block in the block diagonal matrix $C(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\mathrm{MTP}}$ corresponding to the two–factor interactions of the attributes is

$$\begin{aligned}
&\frac{1}{N\Psi_1}\sum_{\boldsymbol{d}}\sum_{a\neq b}\sum_j \psi_a\psi_b c_{\boldsymbol{v}_j} i_{\boldsymbol{v}_j} x_{\boldsymbol{v}_j;\boldsymbol{d},a,b}\left[1-\frac{1}{(1-\ell_{q_1})^{i_{q_1}}(1-\ell_{q_2})^{i_{q_2}}}\right] I_{(\ell_{q_1}-1)(\ell_{q_2}-1)} \\
&= \frac{1}{36\Psi_1}\sum_{a\neq b}\psi_a\psi_b 4\times 1\times 1\left[1-\frac{1}{(1-3)^1(1-3)^1}\right] I_4 \\
&= \frac{\Psi_2}{12\Psi_1} I_4.
\end{aligned}$$

Then we get the $C(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\mathrm{MTP}}$ matrix in Figure 5.3, which has determinant

$$\det(C(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\mathrm{MTP}}) = \frac{(\psi_1\psi_2+\psi_1\psi_3+\psi_2\psi_3)^8}{314928\psi_1\psi_2\psi_3(\psi_1+\psi_2+\psi_3)^{17}}.$$

□

$$C(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\mathrm{MTP}} = \begin{bmatrix} \frac{\Psi_2}{6\Psi_1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{\Psi_2}{6\Psi_1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\Psi_2}{6\Psi_1} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\Psi_2}{6\Psi_1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{\Psi_2}{12\Psi_1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{\Psi_2}{12\Psi_1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{\Psi_2}{12\Psi_1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\Psi_2}{12\Psi_1} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\psi_1+\psi_3}{2\psi_1\psi_3(\psi_1+\psi_2+\psi_3)} & \frac{\psi_1-\psi_3}{2\sqrt{3}\psi_1\psi_3(\psi_1+\psi_2+\psi_3)} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\psi_1-\psi_3}{2\sqrt{3}\psi_1\psi_3(\psi_1+\psi_2+\psi_3)} & \frac{\psi_2\psi_3+\psi_1(\psi_2+4\psi_3)}{6\psi_1\psi_2\psi_3(\psi_1+\psi_2+\psi_3)} \end{bmatrix}$$

Figure 5.3: The information matrix for the design in Example 5.4.4

5.5 A design approach based on complete Latin squares

In this section, we look at a different method of constructing designs for the estimation of contrasts of the entries in $\boldsymbol{\gamma}$ and contrasts of the position effects. The designs that result are not in the class of competing designs of the previous section, since in this section we do not require that all choice sets with difference vector $\boldsymbol{v}_j$ need be in the design if one such choice set is included. This construction method uses the columns of a complete Latin square to give the options in each choice set. An orthogonal array is then used to determine the items that correspond with each of the elements in the complete Latin square. We give some examples of this design, and compare the efficiency of a design constructed in this manner to those discussed in Section 5.4.

Throughout this section we consider only binary experiments, that is

$$\ell_1 = \ell_2 = \ldots = \ell_k = 2.$$

We will begin with an example of a design constructed using this method.

■ **EXAMPLE 5.5.1.**
Suppose that $k = 3$ and $\ell_q = 2$ for $1 \leq q \leq 3$. Then Table 5.6(a) gives a 4 run orthogonal array of strength 2. We also need a 4×4 complete Latin square into which we can embed the orthogonal array. One such complete Latin square is given in the Table 5.6(b).

(a)

0	0	0
0	1	1
1	0	1
1	1	0

(b)

1	2	4	3
2	3	1	4
3	4	2	1
4	1	3	2

Table 5.6: A 4 run orthogonal array of strength 2 (a) and a 4×4 complete Latin square (b).

Then if we replace the 1s in the complete Latin square by the first row of the orthogonal array, the 2s by the second row, and so on, we obtain the design in Table 5.7. □

Option 1	Option 2	Option 3	Option 4
0 0 0	0 1 1	1 1 0	1 0 1
0 1 1	1 0 1	0 0 0	1 1 0
1 0 1	1 1 0	0 1 1	0 0 0
1 1 0	0 0 0	1 0 1	0 1 1

Table 5.7: The 4 choice sets used in Example 5.5.1.

The next theorem shows that this design will lead to a block diagonal information matrix for the estimation of contrasts of the entries in $\boldsymbol{\gamma}$ and contrasts of the position effects.

■ **THEOREM 5.5.1.**

The $C(\boldsymbol{\pi}_0, \boldsymbol{\psi})$ matrix for the estimation of any set of contrasts of the entries in $\boldsymbol{\gamma}$ and any set of contrasts of the entries in $\boldsymbol{\psi}$ will be block diagonal when using a complete Latin square based choice design and the generalised Davidson–Beaver position effects model. □

Proof. Since this design provides a 1–1 mapping of items and distinct entries in the complete Latin square, each item will appear in each position of the ordered choice set exactly once. Therefore $\lambda_{T_i \text{ in pos } a}$ is $\frac{1}{N}$ for all a if T_i appears in the experiment, and 0 if it does not. Then

$$\begin{aligned}
\Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{ia} &= \frac{1}{\Psi_1}\sum_{b\neq a}(\lambda_{T_i \text{ in pos } a} - \lambda_{T_i \text{ in pos } b}) \\
&= \begin{cases} \dfrac{1}{\Psi_1}\displaystyle\sum_{b\neq a}\left(\frac{1}{N}-\frac{1}{N}\right), & \text{if } T_i \text{ appears in the experiment;} \\ \dfrac{1}{\Psi_1}\displaystyle\sum_{b\neq a}(0-0), & \text{otherwise.} \end{cases} \\
&= 0.
\end{aligned}$$

Then $\Lambda_{\gamma\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = \mathbf{0}$, and it follows that

$$C(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = \begin{pmatrix} B_\gamma \Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) B_\gamma^T & \mathbf{0} \\ \mathbf{0} & B_\psi \Lambda_{\psi\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) B_\psi^T \end{pmatrix}.$$

Therefore $C(\boldsymbol{\pi}_0, \boldsymbol{\psi})$ is block diagonal. □

The next example compares the information matrices for the design with 4 choice sets constructed in this manner to a design constructed using the method of Section 5.4.

■ **EXAMPLE 5.5.2.**

The $\Lambda(\boldsymbol{\pi}_0, \boldsymbol{\psi})$ matrix for the design in Table 5.7 is

$$\Lambda(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = \begin{bmatrix} \Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) & \mathbf{0} \\ \mathbf{0} & \Lambda_{\psi\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) \end{bmatrix},$$

where $\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})$ and $\Lambda_{\psi\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi})$ are given in Table 5.8. It follows that the information matrix for the estimation of attribute main effects and contrasts of the position effects is

$$\begin{aligned}
(C(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\mathrm{MP}})_{\mathrm{CLS}} = \text{BlkDiag}\Bigg[&\frac{1}{4(\psi_1+\psi_2+\psi_3+\psi_4)^2} \times \text{Diag}\Bigg[\psi_2\psi_3 + \psi_1\psi_4 + \sum_{a\neq b}\psi_a\psi_b, \\
&2(\psi_2+\psi_3)(\psi_1+\psi_4), \psi_2\psi_3 + \psi_1\psi_4 + \sum_{a\neq b}\psi_a\psi_b\Bigg], C_{\psi\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi})\Bigg],
\end{aligned}$$

where $C_{\psi\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi})$ is shown in Table 5.9. Then the determinant of this information matrix is

$$\det((C(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\mathrm{MP}})_{\mathrm{CLS}}) = \frac{(\psi_2+\psi_3)(\psi_1+\psi_4)(\psi_3\psi_4 + \psi_2(2\psi_3+\psi_4) + \psi_1(\psi_2+\psi_3+2\psi_4))^2}{128\psi_1\psi_2\psi_3\psi_4(\psi_1+\psi_2+\psi_3+\psi_4)^8}.$$

If we compare this to the determinant of the information matrix for an optimal Street–Burgess design as used in Section 5.4, shown in Table 5.10, which is

$$\det((C(\boldsymbol{\pi}_0, \boldsymbol{\psi})_{\mathrm{MP}})_{\text{S–B}}) = \frac{(\psi_1+\psi_2)(\psi_1+\psi_3)(\psi_2+\psi_3)(\psi_1+\psi_4)(\psi_2+\psi_4)(\psi_3+\psi_4)}{32\psi_1\psi_2\psi_3\psi_4(\psi_1+\psi_2+\psi_3+\psi_4)^8},$$

$$\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = \frac{1}{4\,(\psi_1 + \psi_2 + \psi_3 + \psi_4)^{\,2}} \times$$

$$\begin{bmatrix} 2\sum_{a\neq b}\psi_a\psi_b & -\,(\psi_2+\psi_3)\,(\psi_1+\psi_4) & -2\,(\psi_2\psi_3+\psi_1\psi_4) & -\,(\psi_2+\psi_3)\,(\psi_1+\psi_4) \\ -\,(\psi_2+\psi_3)\,(\psi_1+\psi_4) & 2\sum_{a\neq b}\psi_a\psi_b & -\,(\psi_2+\psi_3)\,(\psi_1+\psi_4) & -2\,(\psi_2\psi_3+\psi_1\psi_4) \\ -2\,(\psi_2\psi_3+\psi_1\psi_4) & -\,(\psi_2+\psi_3)\,(\psi_1+\psi_4) & 2\sum_{a\neq b}\psi_a\psi_b & -\,(\psi_2+\psi_3)\,(\psi_1+\psi_4) \\ -\,(\psi_2+\psi_3)\,(\psi_1+\psi_4) & -2\,(\psi_2\psi_3+\psi_1\psi_4) & -\,(\psi_2+\psi_3)\,(\psi_1+\psi_4) & 2\sum_{a\neq b}\psi_a\psi_b \end{bmatrix}$$

$$\Lambda_{\psi\psi}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = \frac{1}{(\psi_1 + \psi_2 + \psi_3 + \psi_4)^{\,2}} \times$$

$$\begin{bmatrix} \frac{\psi_2+\psi_3+\psi_4}{\psi_1} & -1 & -1 & -1 \\ -1 & \frac{\psi_1+\psi_3+\psi_4}{\psi_2} & -1 & -1 \\ -1 & -1 & \frac{\psi_1+\psi_2+\psi_4}{\psi_3} & -1 \\ -1 & -1 & -1 & \frac{\psi_1+\psi_2+\psi_3}{\psi_4} \end{bmatrix}$$

Table 5.8: The block diagonal components of $\Lambda(\boldsymbol{\pi}_0, \boldsymbol{\psi})$ for the complete Latin square based design of Example 5.5.2

$$C_{\psi\psi}(\boldsymbol{\pi}_0,\boldsymbol{\psi}) = \begin{bmatrix} \frac{9\psi_2\psi_3\psi_4+\psi_1(\psi_3\psi_4+\psi_2(9\psi_3+\psi_4))}{20\psi_1\psi_2\psi_3\psi_4(\psi_1+\psi_2+\psi_3+\psi_4)} & \frac{\psi_1(\psi_2(3\psi_3-\psi_4)+\psi_3\psi_4)-3\psi_2\psi_3\psi_4}{4\sqrt{5}\psi_1\psi_2\psi_3\psi_4(\psi_1+\psi_2+\psi_3+\psi_4)} & \frac{3(\psi_2\psi_3\psi_4+\psi_1(\psi_2(\psi_3-\psi_4)-\psi_3\psi_4))}{20\psi_1\psi_2\psi_3\psi_4(\psi_1+\psi_2+\psi_3+\psi_4)} \\ \frac{\psi_1(\psi_2(3\psi_3-\psi_4)+\psi_3\psi_4)-3\psi_2\psi_3\psi_4}{4\sqrt{5}\psi_1\psi_2\psi_3\psi_4(\psi_1+\psi_2+\psi_3+\psi_4)} & \frac{\psi_2\psi_3\psi_4+\psi_1(\psi_3\psi_4+\psi_2(\psi_3+\psi_4))}{4\psi_1\psi_2\psi_3\psi_4(\psi_1+\psi_2+\psi_3+\psi_4)} & \frac{\psi_1(\psi_2(\psi_3+3\psi_4)-3\psi_3\psi_4)-\psi_2\psi_3\psi_4}{4\sqrt{5}\psi_1\psi_2\psi_3\psi_4(\psi_1+\psi_2+\psi_3+\psi_4)} \\ \frac{3(\psi_2\psi_3\psi_4+\psi_1(\psi_2(\psi_3-\psi_4)-\psi_3\psi_4))}{20\psi_1\psi_2\psi_3\psi_4(\psi_1+\psi_2+\psi_3+\psi_4)} & \frac{\psi_1(\psi_2(\psi_3+3\psi_4)-3\psi_3\psi_4)-\psi_2\psi_3\psi_4}{4\sqrt{5}\psi_1\psi_2\psi_3\psi_4(\psi_1+\psi_2+\psi_3+\psi_4)} & \frac{\psi_2\psi_3\psi_4+\psi_1(9\psi_3\psi_4+\psi_2(\psi_3+9\psi_4))}{20\psi_1\psi_2\psi_3\psi_4(\psi_1+\psi_2+\psi_3+\psi_4)} \end{bmatrix}$$

Table 5.9: The $C_{\psi\psi}(\boldsymbol{\pi}_0,\boldsymbol{\psi})$ matrix for the complete Latin square based design of Example 5.5.2

we find that the ratio of the determinants of the information matrices for the two designs can be expressed as

$$\frac{\det((C(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\mathrm{MP}})_{\mathrm{CLS}})}{\det((C(\boldsymbol{\pi}_0,\boldsymbol{\psi})_{\mathrm{MP}})_{\mathrm{S\text{–}B}})} = 1 - \frac{(\psi_2-\psi_3)^2(\psi_1-\psi_4)^2}{(\psi_3\psi_4+\psi_2(2\psi_3+\psi_4)+\psi_1(\psi_2+\psi_3+2\psi_4))^2},$$

which is less than or equal to 1 for all values of $\psi_i \neq 0$. In this case, the design generated using the complete Latin square construction method is at least as efficient as the design in Table 5.10 for the estimation of attribute main effects and contrasts of the position effects when using the Davidson–Beaver position effects model. □

If we construct designs of order 8 using complete Latin squares, we find that the main effects of the attributes cannot always be estimated independently of each other. The following example shows such a situation.

■ **EXAMPLE 5.5.3.**
Suppose that we have an experiment with $k = 5$ and $\ell_i = 2$. We can find an 8 run orthogonal array of strength 2, shown in Table 5.11(a). The defining contrasts for this orthogonal array are $D = AB, E = AC$. We then use the complete Latin square in Table 5.11(b), replacing each entry in the Latin square by the corresponding row of the orthogonal array, to obtain the choice experiment in Table 5.12.

Suppose that we would like to be able to estimate the main effects of the attributes and contrasts of the position effects. Then we can calculate the information matrix, which we find has the form

$$C(\boldsymbol{\pi}_0,\boldsymbol{\psi}) = \begin{bmatrix} a & 0 & 0 & 0 & 0 & \mathbf{0} \\ 0 & b & 0 & 0 & 0 & \mathbf{0} \\ 0 & 0 & c & 0 & 0 & \mathbf{0} \\ 0 & 0 & 0 & d & e & \mathbf{0} \\ 0 & 0 & 0 & e & f & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & C_{\psi\psi}(\boldsymbol{\pi}_0,\boldsymbol{\psi}) \end{bmatrix},$$

where the terms a–f are non–zero. In particular,

$$\begin{aligned} e &= -\psi_1\psi_2+\psi_4\psi_2-\psi_5\psi_2+\psi_8\psi_2+\psi_1\psi_3+\psi_3\psi_4-\psi_3\psi_5-\psi_1\psi_6-\psi_4\psi_6 \\ &\quad +\psi_5\psi_6+\psi_1\psi_7-\psi_4\psi_7+\psi_5\psi_7-\psi_3\psi_8+\psi_6\psi_8-\psi_7\psi_8 \neq 0. \end{aligned}$$

This means that the main effects of the fourth and fifth attributes cannot be estimated independently. □

Option 1			Option 2			Option 3			Option 4		
0	0	0	0	1	1	1	1	0	1	0	1
0	1	1	0	0	0	1	0	1	1	1	0
1	0	1	1	1	0	0	1	1	0	0	0
1	1	0	1	0	1	0	0	0	0	1	1

Table 5.10: The Street–Burgess design used in Example 5.5.2.

(a)

0	0	0	0	0
0	0	1	0	1
0	1	0	1	0
0	1	1	1	1
1	0	0	1	1
1	0	1	1	0
1	1	0	0	1
1	1	1	0	0

(b)

1	2	8	3	7	4	6	5
2	3	1	4	8	5	7	6
8	1	7	2	6	3	5	4
3	4	2	5	1	6	8	7
7	8	6	1	5	2	4	3
4	5	3	6	2	7	1	8
6	7	5	8	4	1	3	2
5	6	4	7	3	8	2	1

Table 5.11: An 8 run orthogonal array of strength 2 (a) and a complete Latin square of order 8 (b).

If we use the 8×8 complete Latin square that has been constructed using the Williams construction, as shown in Table 5.11(b), then we can see which orthogonal arrays of order 8 can be used to estimate the main effects of the attributes independently. Table 5.13 gives the defining relations for each orthogonal array and, where confounds exist, the groups of attributes that are confounded. As stated in Section 1.B, we can construct different complete Latin squares by using different sequencings of groups. Evans [2007] gives a list of such sequencings for groups of small order. When repeating the above analysis with complete Latin squares constructed from each of these different sequencings, we find that the same orthogonal arrays give the same confounded attributes, those shown in Table 5.13.

We finish this section with a discussion about why these confounds occur. Suppose that we let $\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) = [\alpha_{ij}]$, such that $\alpha_{ii} = -\sum_{j \neq i} \alpha_{ij}$. Then we can investigate which contrast matrices, B, give a diagonal $B\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})B^T$ matrix. Let B_U be the main effects contrast matrix for attribute U, B_V be the main effects contrast matrix for attribute V and so on. Now let $B = [B_U B_V B_W B_X B_Y B_Z]^T$, where $B_U = \boldsymbol{j}_2^T \otimes \boldsymbol{j}_2^T \otimes B_2$, $B_V = \boldsymbol{j}_2^T \otimes B_2 \otimes \boldsymbol{j}_2^T$, and $B_W = B_2 \otimes \boldsymbol{j}_2^T \otimes \boldsymbol{j}_2^T$. Given B_U, B_V and B_W, what can we say about B_X, B_Y and B_Z?

The orthogonality of the contrast matrix imposes the constraints

$$\begin{aligned} B_U B_X^T &= 0, \\ B_V B_X^T &= 0, \end{aligned}$$

and

$$B_W B_X^T = 0,$$

on the entries of B_X. Since the entries in B_X are contrast coefficients, we also require

$$B_X \boldsymbol{j}^T = 0.$$

In this example, we would like $C_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})$ to be diagonal, which places further constraints on the entries in B_X. These constraints are

$$\begin{aligned} B_U \Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) B_X^T &= 0, \\ B_V \Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) B_X^T &= 0, \end{aligned}$$

Option 1	Option 2	Option 3	Option 4	Option 5	Option 6	Option 7	Option 8
0 0 0 0 0	0 0 1 0 1	1 1 1 0 0	0 1 0 1 0	1 1 0 0 1	0 1 1 1 1	1 0 1 1 0	1 0 0 1 1
0 0 1 0 1	0 1 0 1 0	0 0 0 0 0	0 1 1 1 1	1 1 1 0 0	1 0 0 1 1	1 1 0 0 1	1 0 1 1 0
1 1 1 0 0	0 0 0 0 0	1 1 0 0 1	0 0 1 0 1	1 0 1 1 0	0 1 0 1 0	1 0 0 1 1	0 1 1 1 1
0 1 0 1 0	0 1 1 1 1	0 0 1 0 1	1 0 0 1 1	0 0 0 0 0	1 0 1 1 0	1 1 1 0 0	1 1 0 0 1
1 1 0 0 1	1 1 1 0 0	1 0 1 1 0	0 0 0 0 0	1 0 0 1 1	0 0 1 0 1	0 1 1 1 1	0 1 0 1 0
0 1 1 1 1	1 0 0 1 1	0 1 0 1 0	1 0 1 1 0	0 0 1 0 1	1 1 0 0 1	0 0 0 0 0	1 1 1 0 0
1 0 1 1 0	1 1 0 0 1	1 0 0 1 1	1 1 1 0 0	0 1 1 1 1	0 0 0 0 0	0 1 0 1 0	0 0 1 0 1
1 0 0 1 1	1 0 1 1 0	0 1 1 1 1	1 1 0 0 1	0 1 0 1 0	1 1 1 0 0	0 0 1 0 1	0 0 0 0 0

Table 5.12: Complete Latin square based design used in Example 5.5.3.

k	Defining Relation	Confounded Attributes
3	None	None
4	D=ABC	(1,4)
4	D=AB	None
5	D=AB, E=AC	None
5	D=AB, E=BC	(4,5)
5	D=AC. E=BC	None
6	D=AB, E=AC, F=BC	(4,5)
7	D=AB, E=AC, F=BC, G=ABC	(1,7) and (4,5)

Table 5.13: Confounded attributes when a complete Latin square based design of order 8 is used

and

$$B_W \Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) B_X^T \quad = \quad 0.$$

It should be noted at this point that B_U, B_V, and B_W have been constructed in such a way that they are orthogonal to each other, and produce 0 entries in the relevant off–diagonal entries in $C_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})$. Suppose that B_X is given by

$$B_X = [x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8],$$

and $\Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})$ is given by

$$\begin{bmatrix} D & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_3 & \alpha_2 & \alpha_1 \\ \alpha_1 & D & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_3 & \alpha_2 \\ \alpha_2 & \alpha_1 & D & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_3 \\ \alpha_3 & \alpha_2 & \alpha_1 & D & \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \\ \alpha_4 & \alpha_3 & \alpha_2 & \alpha_1 & D & \alpha_1 & \alpha_2 & \alpha_3 \\ \alpha_3 & \alpha_4 & \alpha_3 & \alpha_2 & \alpha_1 & D & \alpha_1 & \alpha_2 \\ \alpha_2 & \alpha_3 & \alpha_4 & \alpha_3 & \alpha_2 & \alpha_1 & D & \alpha_1 \\ \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 & \alpha_3 & \alpha_2 & \alpha_1 & D \end{bmatrix},$$

where

$$\begin{aligned} \alpha_1 &= \frac{\psi_1\psi_2 + \psi_1\psi_3 + \psi_2\psi_4 + \psi_3\psi_5 + \psi_4\psi_6 + \psi_5\psi_7 + \psi_6\psi_8 + \psi_7\psi_8}{8\Psi_1}, \\ \alpha_2 &= \frac{\psi_1\psi_6 + \psi_1\psi_7 + \psi_2\psi_5 + \psi_2\psi_8 + \psi_3\psi_4 + \psi_3\psi_8 + \psi_4\psi_7 + \psi_5\psi_6}{8\Psi_1}, \\ \alpha_3 &= \frac{\psi_1\psi_4 + \psi_1\psi_5 + \psi_2\psi_3 + \psi_2\psi_6 + \psi_3\psi_7 + \psi_4\psi_8 + \psi_5\psi_8 + \psi_6\psi_7}{8\Psi_1}, \end{aligned}$$

and

$$\alpha_4 \quad = \quad \frac{\psi_1\psi_8 + \psi_2\psi_7 + \psi_3\psi_6 + \psi_4\psi_5}{4\Psi_1}.$$

Then we can expand $B_U \Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) B_X^T$ to obtain

$$\begin{aligned} B_U \Lambda_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi}) B_X^T \;=\; & 2\alpha_1(-x_1 - x_4 + x_5 + x_8) + 2\alpha_2(-x_1 - x_2 - x_3 - x_4 + x_5 + x_6 + x_7 + x_8) \\ & +2\alpha_3(-x_1 - 2x_2 - 2x_3 - x_4 + x_5 + 2x_6 + 2x_7 + x_8) \\ & +2\alpha_4(-x_1 - x_2 - x_3 - x_4 + x_5 + x_6 + x_7 + x_8), \end{aligned}$$

which we require to be equal to 0 for all values of α_i. By equating coefficients of α_i, we find

$$\begin{aligned} -x_1 - x_4 + x_5 + x_8 &= 0, \\ -x_1 - x_2 - x_3 - x_4 + x_5 + x_6 + x_7 + x_8 &= 0, \end{aligned}$$

and

$$-x_1 - 2x_2 - 2x_3 - x_4 + x_5 + 2x_6 + 2x_7 + x_8 = 0.$$

We also have the constraints

$$\begin{aligned} -x_1 - x_2 - x_3 - x_4 + x_5 + x_6 + x_7 + x_8 &= 0, \\ -x_1 - x_2 + x_3 + x_4 - x_5 - x_6 + x_7 + x_8 &= 0, \end{aligned}$$

and

$$-x_1 + x_2 - x_3 + x_4 - x_5 + x_6 - x_7 + x_8 = 0,$$

that make the contrasts orthogonal to each other, and

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 = 0,$$

to make B_X a set of contrast coefficients. These reduce to give the following 5 linearly independent equations

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 &= 0, \\ -x_1 - x_2 - x_3 - x_4 + x_5 + x_6 + x_7 + x_8 &= 0, \\ -x_1 - x_2 + x_3 + x_4 - x_5 - x_6 + x_7 + x_8 &= 0, \\ -x_1 + x_2 - x_3 + x_4 - x_5 + x_6 - x_7 + x_8 &= 0, \end{aligned}$$

and

$$-x_1 - x_4 + x_5 + x_8 = 0.$$

The general solution to these equations is given by

$$x_1 = x_8, \quad x_2 = x_7, \quad x_3 = x_6, \quad x_4 = x_5, \quad x_4 = -x_6 - x_7 - x_8.$$

One possible solution is

$$B_X = \frac{1}{2\sqrt{2}}[-1, -1, 1, 1, 1, 1, -1, -1].$$

If we use this contrast, we can construct a fifth contrast to be estimated in the same manner as B_X. Suppose that this new contrast has coefficients

$$B_Y = [y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_8].$$

The constraints on the y_i will include all of those that were imposed on the x_i, plus some more to ensure that B_X and B_Y are orthogonal, and have a 0 in the corresponding entry of $C_{\gamma\gamma}(\boldsymbol{\pi}_0, \boldsymbol{\psi})$. By repeating the process that we used to derive B_X, we have the following seven linearly independent constraints on the y_i:

$$\begin{aligned}
y_1+y_2+y_3+y_4+y_5+y_6+y_7+y_8 &= 0,\\
-y_1-y_2-y_3-y_4+y_5+y_6+y_7+y_8 &= 0,\\
-y_1-y_2+y_3+y_4-y_5-y_6+y_7+y_8 &= 0,\\
-y_1+y_2-y_3+y_4-y_5+y_6-y_7+y_8 &= 0,\\
-y_1-y_4+y_5+y_8 &= 0,\\
-y_1-y_2+y_3+y_4+y_5+y_6-y_7-y_8 &= 0,
\end{aligned}$$

and

$$-y_1+y_4+y_5-y_8 = 0.$$

This gives the general solution

$$y_1=y_4=y_5=y_8, \qquad y_2=y_3=y_6=y_7.$$

One possible solution is

$$B_Y=\frac{1}{2\sqrt{2}}[-1,1,1,-1,-1,1,1,-1].$$

If we continue this process, we will have one additional constraint on B_Z that is independent of the existing constraints. This constraint is

$$-z_1+z_2+z_3-z_4-z_5+z_6+z_7-z_8=0.$$

The only consistent set of solutions to these constraints is $z_i=0$ for $1\le i\le 8$. Therefore, no additional contrasts can be estimated independently without forming a non–diagonal $C_{\gamma\gamma}(\boldsymbol{\pi}_0,\boldsymbol{\psi})$ matrix.

Thus, we can see that while designs constructed using complete Latin squares can be more efficient than those designs used in Section 5.4 in some cases, there are restrictions on the number of attributes that we can use in designs.

5.6 Simulations of the generalised Davidson–Beaver position effects model

In this section we consider the performance of the generalised Davidson–Beaver position effects model under various model assumptions by carrying out a number of simulation studies. We assume that $k=2$, $\ell_1=\ell_2=2$ and $m=3$ throughout. We consider two sets of values for the parameters. In the first we assume that both main effects parameters, τ_1 and τ_2, are equal to 0 and the position main effect parameters are equal to $\psi_{\text{L}}=-0.3$ and $\psi_{\text{Q}}=0.1$, and in the second set we assume that $\tau_1=1$ and $\tau_2=-1$ but $\psi_{\text{L}}=-0.3$ and $\psi_{\text{Q}}=0.1$ still.

We find efficient designs for each set of values and compare the performance of each design with both sets of parameter values. The design in Table 5.14 is optimal for the estimation of the main effects of the attributes plus the position main effects when $\tau_1=\tau_2=0,\psi_{\text{L}}=-0.3$, and $\psi_{\text{Q}}=0.1$, by Theorem 5.4.7. An alternative design is shown in Table 5.15, which is more efficient than the design in Table 5.14 when $\tau_1=1,\tau_2=-1,\psi_{\text{L}}=-0.3$, and $\psi_{\text{Q}}=0.1$.

Option 1	Option 2	Option 3	P(T_1)	P(T_2)	P(T_3)
0 0	0 1	1 0	0.195	0.014	0.791
0 1	0 0	1 1	0.110	0.445	0.445
1 0	1 1	0 0	0.871	0.065	0.065
1 1	1 0	0 1	0.195	0.791	0.014

Table 5.14: Optimal design for the estimation of attribute main effects and contrasts of the position effects when $\tau_1 = \tau_2 = 0, \psi_L = -0.3$, and $\psi_Q = 0.1$, with selection probabilities when $\tau_1 = 1, \tau_2 = -1, \psi_L = -0.3$, and $\psi_Q = 0.1$.

Option 1	Option 2	Option 3
0 0	0 1	1 0
0 1	0 0	1 1
1 0	1 1	0 0
1 0	1 1	0 1
1 1	1 0	0 1

Table 5.15: A design more efficient than the design in Table 5.14 when $\tau_1 = 1, \tau_2 = -1, \psi_L = -0.3$, and $\psi_Q = 0.1$.

We first assume that $\tau_1 = \tau_2 = 0, \psi_L = -0.3$, and $\psi_Q = 0.1$ and compare the simulated distributions of the parameter estimates when the designs in Tables 5.14 and 5.15 are used in turn. Each simulation is modelled using the simulated responses from 150 respondents, and each boxplot displays the distribution of the estimates from 1000 such simulations. Figures 5.4(a) and (b) show the distributions of the parameter estimates when the designs in Tables 5.14 and 5.15, respectively, are used. Summary statistics for both simulations are provided in Table 5.16. We see that, for both designs, the distribution of the parameter estimates seem to be unbiased and symmetric. We see that, in this case, the additional choice set in the design in Table 5.15 does not seem to reduce the variance of the parameter estimates.

We now consider the performance of these two designs when $\tau_1 = 1, \tau_2 = -1, \psi_L = -0.3$, and $\psi_Q = 0.1$. Figures 5.5(a) and (b) show the distributions of the parameter estimates when the designs in Tables 5.14 and 5.15, respectively, are used. Summary statistics for both simulations are provided in Table 5.17. We see that, for both designs, the distribution of the parameter estimates seem to be unbiased and close to symmetric. For these parameter estimates, we see that the addition of an extra choice set does seem to reduce the variance of the parameter estimates. The selection probabilities when $\tau_1 = 1, \tau_2 = -1, \psi_L = -0.3$, and $\psi_Q = 0.1$ for the design in Table 5.14 are given in the last three columns of Table 5.14.

Next, we compare the distributions of the parameter estimates for different values of ψ_L and ψ_Q when the design in Table 5.15 is used. Suppose that $\tau_1 = 0.5$, and $\tau_2 = -1$. Then Figures 5.6(a) and (b) show the distributions of the parameter estimates when the coefficient of the linear component of the position effect is -0.2 and -0.4, respectively, with a zero quadratic component

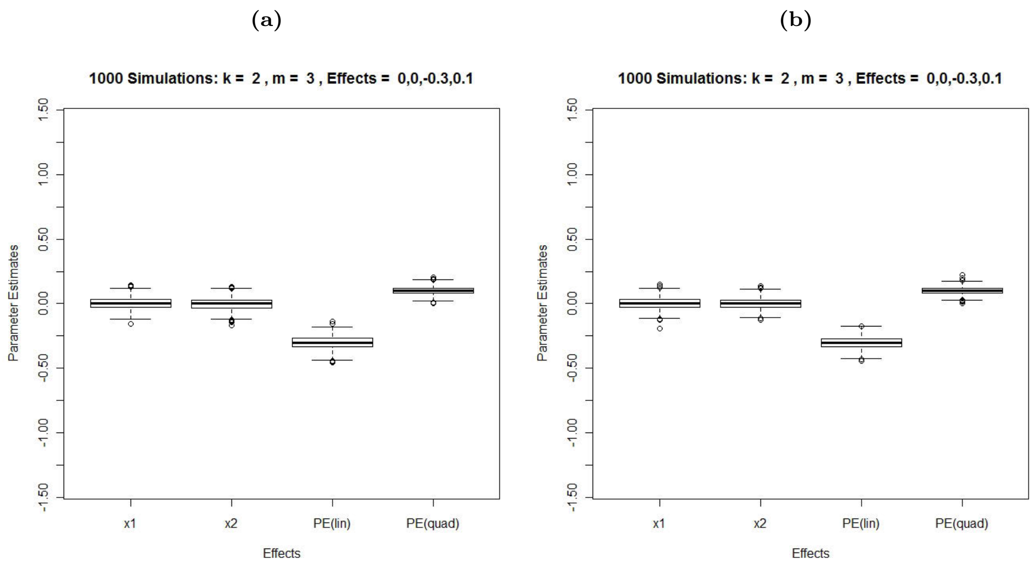

Figure 5.4: Simulations: $\tau_1 = \tau_2 = 0$, $\psi_{\mathrm{L}} = -0.3$, and $\psi_{\mathrm{Q}} = 0.1$.

Parameter	Simulated Mean (Standard Error)	Theoretical Variance	Simulated Variance	Simulated Skewness (Standard Error)
Design in Table 5.14				
τ_1	0.00065(0.00145)	0.00203	0.00209	0.09595(0.07734)
τ_2	−0.00244(0.00147)	0.00191	0.00217	−0.03946(0.07734)
ψ_{L}	−0.30190(0.00159)	0.00748	0.00251	−0.14780(0.07734)
ψ_{Q}	0.10039(0.00099)	0.00254	0.00098	0.08010(0.07734)
Design in Table 5.15				
τ_1	0.00074(0.00138)	0.00184	0.00191	−0.01432(0.07734)
τ_2	0.00060(0.00134)	0.00172	0.00178	−0.03239(0.07734)
ψ_{L}	−0.30178(0.00144)	0.00719	0.00206	0.07623(0.07734)
ψ_{Q}	0.10125(0.00090)	0.00218	0.00080	0.08965(0.07734)

Table 5.16: Summary statistics for $\tau_1 = \tau_2 = 0$, $\psi_{\mathrm{L}} = -0.3$, and $\psi_{\mathrm{Q}} = 0.1$.

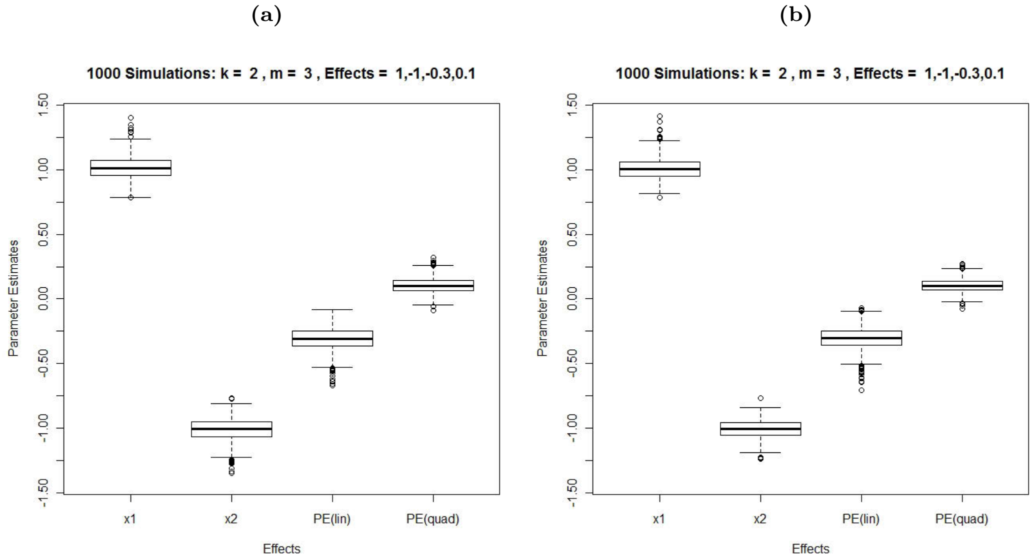

Figure 5.5: Simulations: $\tau_1 = 1$, $\tau_2 = -1$, $\psi_L = -0.3$, and $\psi_Q = 0.1$.

Parameter	**Simulated Mean (Standard Error)**	**Theoretical Variance**	**Simulated Variance**	**Simulated Skewness (Standard Error)**
Design in Table 5.14				
τ_1	1.01674(0.00274)	0.00245	0.00749	0.46021(0.07734)
τ_2	−1.01164(0.00275)	0.00221	0.00758	−0.39469(0.07734)
ψ_L	−0.31189(0.00283)	0.00970	0.00800	−0.40517(0.07734)
ψ_Q	0.10227(0.00188)	0.00292	0.00354	0.23414(0.07734)
Design in Table 5.15				
τ_1	1.00935(0.00263)	0.00184	0.00693	0.51804(0.07734)
τ_2	−1.00761(0.00212)	0.00173	0.00450	−0.18823(0.07734)
ψ_L	−0.30817(0.00273)	0.00751	0.00748	−0.52933(0.07734)
ψ_Q	0.10126(0.00159)	0.00209	0.00254	0.00055(0.07734)

Table 5.17: Summary statistics for $\tau_1 = 1$, $\tau_2 = -1$, $\psi_L = -0.3$, and $\psi_Q = 0.1$.

in both cases. Figures 5.7 (a) and (b) show the distributions of the parameter effects when the coefficient of the linear component of the position effect is assumed to be -0.2 and -0.4, respectively, and the coefficient of the quadratic component of the position effect is 0.2 in both cases. Summary statistics for all four simulations are provided in Table 5.18. For each of the sets of parameter values, this design gives unbiased and close to symmetric parameter estimates. The variance of these parameter estimates remains relatively constant across the values of ψ_{L} and ψ_{Q}, demonstrating the robustness of this design.

We now look at the ability of a range of designs to estimate the main effects plus two–factor interactions of the attributes and the position main effects. The first two designs we consider are those in Tables 5.14 and 5.15. The third design is that in Table 5.19, which is optimal for the estimation of the main effects plus the two–factor interaction when the MNL model is used and $\tau_1 = \tau_2 = \tau_{12} = 0$, by Theorem 1.3.5. The final design is shown in Table 5.20, and is more efficient than the design in Table 5.19 for estimating the main effects plus the two–factor interaction of the attributes and the position main effects when $\tau_1 = 1, \tau_2 = -1, \tau_{12} = -0.25, \psi_{\mathrm{L}} = -0.3$, and $\psi_{\mathrm{Q}} = 0.1$.

First, we consider the case where the interaction is assumed to be negligible. We let the coefficients of the attribute main effects be $\tau_1 = 1$, and $\tau_2 = -1$ as before, with the coefficient of the linear component of the main effect fixed to be -0.3 and the coefficient of the quadratic component fixed to be 0.1. Then Figures 5.8(a), (b), (c), and (d) give simulated distributions when the designs in Table 5.14, Table 5.15, Table 5.19, and Table 5.20 are used. Summary statistics for all four of the simulations are provided in Table 5.21.

Both the design in Table 5.19 and the design in Table 5.20 give unbiased and symmetric parameter estimates with relatively small variances. Slightly more bias and skewness can be seen in the estimates from the designs in Tables 5.14 and 5.15, as well as larger variances. The larger variances are not surprising, since the designs in Tables 5.19 and 5.20 contain about three times as many choice sets as the designs in Tables 5.14 and 5.15.

Now suppose that we have a non–zero interaction between the attributes. We let the coefficients of the attribute main effects and the contrasts of the position effects be the same as the zero–interaction case, and fix $\tau_{12} = -0.25$. Figures 5.9(a), (b), (c), and (d) give simulated distributions when the designs in Table 5.14, Table 5.15, Table 5.19, and Table 5.20 are used. Summary statistics for all four of the simulations are provided in Table 5.22.

Again, both the design in Table 5.19 and the design in Table 5.20 give unbiased and symmetric parameter estimates with relatively small variances. Once again, we see slightly more bias and skewness in the estimates from the designs in Tables 5.14 and 5.15, as well as larger variances.

(a) **(b)**

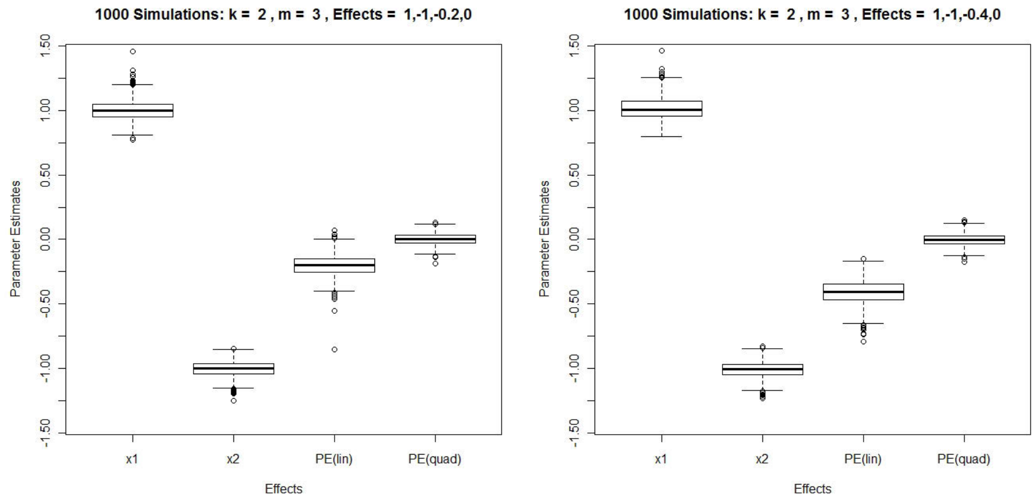

Figure 5.6: Simulations: $\tau_1 = 1$, $\tau_2 = -1$, (a) $\psi_{\rm L} = -0.2$ and (b) $\psi_{\rm L} = -0.4$, and $\psi_{\rm Q} = 0$.

(a) **(b)**

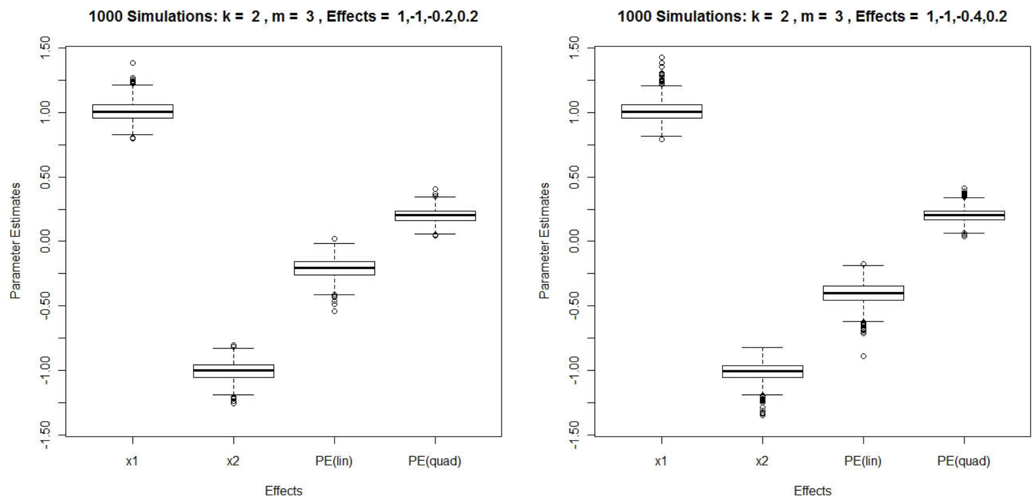

Figure 5.7: Simulations: $\tau_1 = 1$, $\tau_2 = -1$, (a) $\psi_{\rm L} = -0.2$ and (b) $\psi_{\rm L} = -0.4$, and $\psi_{\rm Q} = 0.2$.

Parameter	Simulated Mean (Standard Error)	Theoretical Variance	Simulated Variance	Simulated Skewness (Standard Error)
$\tau_1 = 1$, $\tau_2 = -1$,$\psi_L = -0.2$, and $\psi_Q = 0$				
τ_1	1.00573(0.00259)	0.00179	0.00671	1.15483(0.07734)
τ_2	−1.00355(0.00190)	0.00172	0.00359	−0.27256(0.07734)
ψ_L	−0.20135(0.00264)	0.00766	0.00696	−0.67568(0.07734)
ψ_Q	0.00031(0.00138)	0.00206	0.00190	−0.09086(0.07734)
$\tau_1 = 1$, $\tau_2 = -1$,$\psi_L = -0.4$, and $\psi_Q = 0$				
τ_1	1.01584(0.00287)	0.00194	0.00825	0.45435(0.07734)
τ_2	−1.00950(0.00199)	0.00170	0.00394	−0.21212(0.07734)
ψ_L	−0.41253(0.00298)	0.00768	0.00887	−0.34218(0.07734)
ψ_Q	−0.00264(0.00149)	0.00204	0.00222	0.00867(0.07734)
$\tau_1 = 1$, $\tau_2 = -1$,$\psi_L = -0.2$, and $\psi_Q = 0.2$				
τ_1	1.01168(0.00243)	0.00176	0.00592	0.38612(0.07734)
τ_2	−1.00968(0.00224)	0.00176	0.00504	−0.34004(0.07734)
ψ_L	−0.21025(0.00250)	0.00739	0.00623	−0.44861(0.07734)
ψ_Q	0.20143(0.00167)	0.00214	0.00280	0.15214(0.07734)
$\tau_1 = 1$, $\tau_2 = -1$,$\psi_L = -0.4$, and $\psi_Q = 0.2$				
τ_1	1.01265(0.00273)	0.00189	0.00744	0.63841(0.07734)
τ_2	−1.01266(0.00237)	0.00174	0.00564	−0.63725(0.07734)
ψ_L	−0.40498(0.00277)	0.00741	0.00768	−0.55272(0.07734)
ψ_Q	0.20355(0.00170)	0.00212	0.00290	0.20674(0.07734)

Table 5.18: Summary statistics for $\tau_1 = 1, \tau_2 = -1$ and various values for ψ_L and ψ_Q.

Option 1	Option 2	Option 3	Option 1	Option 2	Option 3
0 0	0 1	1 0	1 0	0 0	0 1
0 1	0 0	1 1	1 1	0 1	0 0
1 0	1 1	0 0	0 0	0 1	1 1
1 1	1 0	0 1	0 1	0 0	1 0
0 0	1 0	1 1	1 0	1 1	0 1
0 1	1 1	1 0	1 1	1 0	0 0

Table 5.19: Optimal design for the estimation of attribute main effects and two–factor interactions when $\tau_1 = \tau_2 = \tau_{12} = 0, \psi_L = -0.3$, and $\psi_Q = 0.1$.

Option 1	Option 2	Option 3	Option 1	Option 2	Option 3
0 0	0 1	1 1	1 0	0 0	0 1
0 1	0 0	1 1	1 1	0 1	0 0
1 0	1 1	0 0	0 1	0 0	1 0
1 1	1 0	0 1	1 0	1 1	0 1
0 0	1 0	1 1	1 1	1 0	0 0
0 1	1 1	1 0			

Table 5.20: Optimal design for the estimation of attribute main effects and two–factor interactions when $\tau_1 = 1, \tau_2 = -1, \tau_{12} = -0.25, \psi_L = -0.3$, and $\psi_Q = 0.1$.

Parameter	Simulated Mean (Standard Error)	Theoretical Variance	Simulated Variance	Simulated Skewness (Standard Error)
Design in Table 5.14				
τ_1	1.01505(0.00301)	0.00251	0.00907	0.41817(0.07734)
τ_2	−1.01785(0.00306)	0.00226	0.00934	−0.45840(0.07734)
τ_{12}	−0.00464(0.00224)	0.00192	0.00503	−0.21515(0.07734)
ψ_L	−0.31117(0.00281)	0.00970	0.00789	−0.57501(0.07734)
ψ_Q	0.10462(0.00192)	0.00292	0.00369	0.37087(0.07734)
Design in Table 5.15				
τ_1	1.01857(0.00294)	0.00184	0.00865	0.55270(0.07734)
τ_2	−1.01279(0.00247)	0.00185	0.00609	−0.26282(0.07734)
τ_{12}	−0.00673(0.00220)	0.00180	0.00486	−0.18801(0.07734)
ψ_L	−0.30781(0.00275)	0.00752	0.00756	−0.47014(0.07734)
ψ_Q	0.09889(0.00157)	0.00233	0.00245	−0.01713(0.07734)
Design in Table 5.19				
τ_1	1.00307(0.00146)	0.00073	0.00215	0.13036(0.07734)
τ_2	−1.00299(0.00140)	0.00069	0.00195	−0.03680(0.07734)
τ_{12}	−0.00155(0.00142)	0.00068	0.00202	−0.07224(0.07734)
ψ_L	−0.30201(0.00130)	0.00288	0.00170	−0.04205(0.07734)
ψ_Q	0.10111(0.00077)	0.00090	0.00059	0.07402(0.07734)
Design in Table 5.20				
τ_1	1.00217(0.00157)	0.00077	0.00246	0.12641(0.07734)
τ_2	−1.00306(0.00142)	0.00076	0.00203	−0.13142(0.07734)
τ_{12}	−0.00169(0.00140)	0.00077	0.00196	−0.11987(0.07734)
ψ_L	−0.29905(0.00139)	0.00308	0.00194	−0.00880(0.07734)
ψ_Q	0.10002(0.00077)	0.00101	0.00059	−0.03074(0.07734)

Table 5.21: Summary statistics for $\tau_1 = 1, \tau_2 = -1, \tau_{12} = 0, \psi_L = -0.3$ and $\psi_Q = 0.1$.

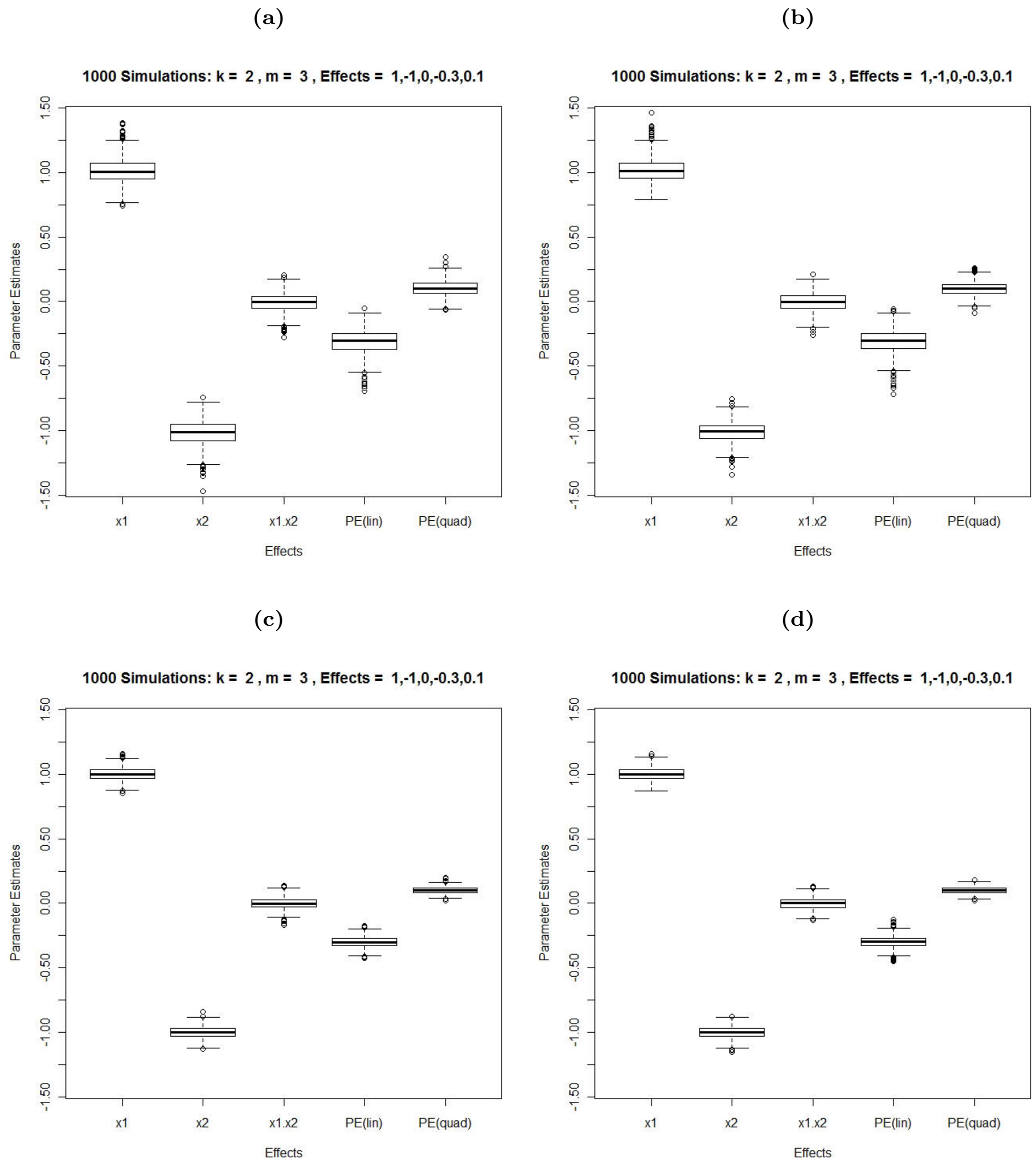

Figure 5.8: Simulations: Estimating attribute main effects and contrasts of the position effects, designs in (a) Table 5.14, (b) Table 5.15, (c) Table 5.19, and (d) Table 5.20.

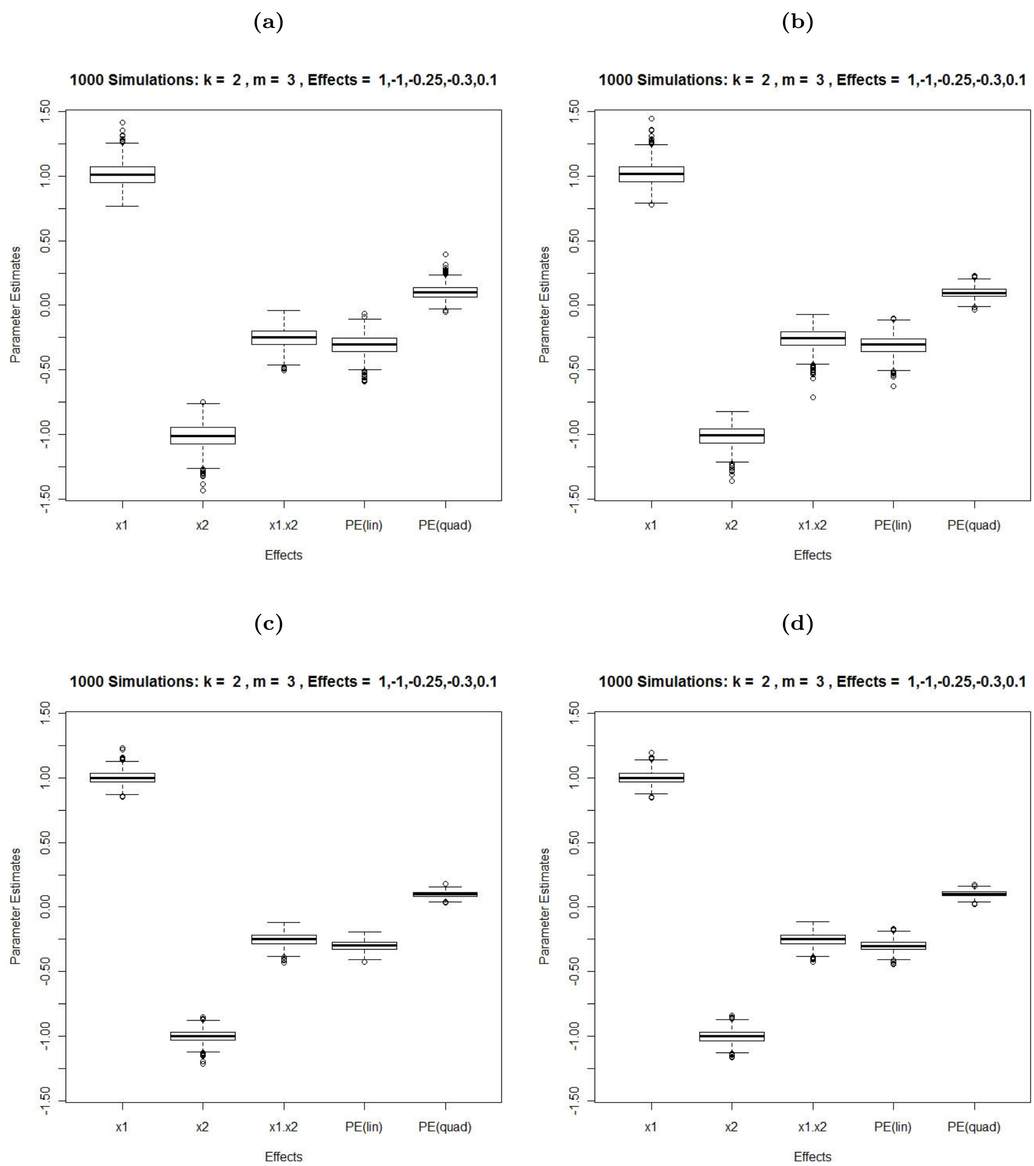

Figure 5.9: Simulations: Estimating attribute main effects, two–factor interactions and contrasts of the position effects, designs in (a) Table 5.14, (b) Table 5.15, (c) Table 5.19, and (d) Table 5.20.

Parameter	Simulated Mean (Standard Error)	Theoretical Variance	Simulated Variance	Simulated Skewness (Standard Error)
Design in Table 5.14				
τ_1	1.01295(0.00298)	0.00250	0.00887	0.33390(0.07734)
τ_2	−1.01650(0.00301)	0.00225	0.00906	−0.44446(0.07734)
τ_{12}	−0.25309(0.00252)	0.00193	0.00633	−0.31484(0.07734)
ψ_{L}	−0.30672(0.00244)	0.00999	0.00597	−0.32199(0.07734)
ψ_{Q}	0.10299(0.00171)	0.00286	0.00291	0.52240(0.07734)
Design in Table 5.15				
τ_1	1.02163(0.00288)	0.00184	0.00829	0.34026(0.07734)
τ_2	−1.01232(0.00250)	0.00186	0.00624	−0.45037(0.07734)
τ_{12}	−0.26029(0.00247)	0.00179	0.00613	−0.58993(0.07734)
ψ_{L}	−0.30944(0.00239)	0.00762	0.00574	−0.28540(0.07734)
ψ_{Q}	0.09738(0.00135)	0.00229	0.00182	0.02016(0.07734)
Design in Table 5.20				
τ_1	1.00222(0.00162)	0.00073	0.00263	0.36760(0.07734)
τ_2	−1.00286(0.00161)	0.00069	0.00259	−0.29557(0.07734)
τ_{12}	−0.24988(0.00155)	0.00068	0.00240	−0.27733(0.07734)
ψ_{L}	−0.30057(0.00122)	0.00287	0.00150	−0.14386(0.07734)
ψ_{Q}	0.09990(0.00070)	0.00090	0.00049	0.02138(0.07734)
Design in Table 5.19				
τ_1	1.00448(0.00168)	0.00078	0.00283	0.27812(0.07734)
τ_2	−1.00390(0.00167)	0.00077	0.00280	−0.22667(0.07734)
τ_{12}	−1.00390(0.00167)	0.00077	0.00280	−0.22667(0.07734)
ψ_{L}	−0.30149(0.00132)	0.00305	0.00174	−0.08309(0.07734)
ψ_{Q}	0.10114(0.00069)	0.00101	0.00048	0.02933(0.07734)

Table 5.22: Summary statistics for $\tau_1 = 1, \tau_2 = -1, \tau_{12} = -0.25, \psi_{\mathrm{L}} = -0.3$ and $\psi_{\mathrm{Q}} = 0.1$.

5.A Proof that $\mathbf{j}_\mathbf{L}^\mathbf{T}\mathbf{z} = 0$ for the generalised Davidson–Beaver Position Effects Model

We begin by recalling that

$$z_i = \sum_{i \in C} w_{i|C} - \frac{sn_C \widehat{\psi_a} \widehat{\pi_i}}{\sum_{b=1}^m \widehat{\psi_b} \widehat{\pi_{i_b}}},$$

Now, the vector $\boldsymbol{z}$ contains the values for z_i for each possible item T_i. Then

$$\begin{aligned}
\boldsymbol{j}_L^T \boldsymbol{z} &= \sum_{i=1}^t z_i \\
&= \sum_i \sum_{i \in C} \left(w_{i|C} - \frac{sn_C \widehat{\psi_a} \widehat{\pi_i}}{\sum_{b=1}^m \widehat{\psi_b} \widehat{\pi_{i_b}}} \right) \\
&= \sum_i \sum_{i \in C} (w_{i|C}) - \sum_C \frac{sn_C \sum_{b=1}^m \widehat{\psi_b} \widehat{\pi_{i_b}}}{\sum_{b=1}^m \widehat{\psi_b} \widehat{\pi_{i_b}}} \\
&= \sum_C sn_C - \sum_C sn_C \\
&= 0,
\end{aligned}$$

as required.

Chapter 6

Optimal Designs when using Fractional Factorial Starting Designs

In Section 1.3 we gave results about the construction of optimal choice experiments which were obtained from a starting design to which we added a suitable set of generators. In all of the results so far, the starting design has been a full factorial design. In this chapter, we will give some results about the construction of choice experiments where the starting design is a fractional factorial design. We will restrict our discussion to symmetric designs with a prime power number of levels.

We begin this chapter by looking at the construction of contrast matrices that contain the contrast coefficients of only the items in that fraction, when main effects are of interest. We then consider the construction of contrast matrices when a generator is added to the starting design, where the addition is performed component–wise in $GF[\ell]$. We work in $GF[\ell]$ to take advantage of the field properties, in particular the fact that $xy = 0$ implies that at least one of x and y is 0 (so $2 \times 3 = 0 \mod 6$ but neither 2 nor 3 is 0). Finally, we prove a theorem that gives rules for the optimal design for the estimation of main effects by using the properties of these contrast matrices and of the associated choice sets.

6.1 Constructing a Contrast Matrix for Regular Designs

Consider an ℓ^{k-p} regular fractional factorial starting design, where ℓ is a prime or a prime power. We can reorder the rows and columns of the design so the first $k-p$ columns form a complete ℓ^{k-p} factorial with the rows in lexicographic order. Then, the rows of the contrast matrix corresponding to the first $k-p$ attributes are given by

$$B_{F,(1,2,\ldots,k-p)} = \begin{bmatrix} B_\ell \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes \ldots \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \\ \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes B_\ell \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes \ldots \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \\ \vdots \\ \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes \ldots \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes B_\ell \end{bmatrix},$$

where B_ℓ is an $(\ell-1)\times\ell$ matrix of orthogonal polynomial contrast coefficients with $B_\ell B_\ell^T = I_{\ell-1}$, and $\boldsymbol{j}_\ell$ is an $\ell \times 1$ column vector of 1s, as defined in Section 1.2.

We use a slightly altered version of the Rao–Hamming construction to obtain the remaining attributes. Instead of the usual constraint, that the first non–zero coefficient to be equal to 1, we constrain the last non–zero coefficient to be equal to 1. This alteration will simplify the description of the contrast matrices later on. In this chapter we will restrict our scope to designs constructed in this way.

If we recall Construction 1.B.1, we note that the construction begins with an ℓ^{k-p} full factorial design in the columns C_1 to C_{k-p} of the starting design and we obtain the remaining p columns of the starting design from linear combinations of the first $k-p$ columns, with addition conducted component–wise in $GF[\ell]$. Thus we let $\boldsymbol{b}_1$ be a row vector that contains the levels for the first attribute, $\boldsymbol{b}_2$ be a row vector that contains the levels for the second attribute, and so forth up to $\boldsymbol{b}_k$ be a row vector that contains the levels for the k^{th} attribute, and we define $\boldsymbol{b}_{k-p+1}, \ldots, \boldsymbol{b}_k$ as linear combinations of $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{k-p}$. Note that by our assumption, the last non–zero $\boldsymbol{b}_i$ will have a coefficient of 1. This ensures that we produce a set of attributes that form a fractional factorial design. We now consider two examples of this construction.

■ **EXAMPLE 6.1.1.**

In this example, we construct a 3^{4-2} regular fractional factorial starting design in 9 runs. We begin with two attributes that form a 3^2 complete factorial design. We define two further attributes as linear combinations of the first two attributes. We define these linear combinations by considering all linear combinations of $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ where the coefficient of $\boldsymbol{b}_2$ is equal to 1. The additions carried out in this construction use the addition rules of $GF[3]$. This gives

$$\begin{aligned}
\boldsymbol{b}_1 &= (0,0,0,1,1,1,2,2,2),\\
\boldsymbol{b}_2 &= (0,1,2,0,1,2,0,1,2),\\
\boldsymbol{b}_3 &= \boldsymbol{b}_1 + \boldsymbol{b}_2\\
&= (0,0,0,1,1,1,2,2,2) + (0,1,2,0,1,2,0,1,2)\\
&= (0,1,2,1,2,0,2,0,1),
\end{aligned}$$

and

$$\begin{aligned}
\boldsymbol{b}_4 &= 2 \times \boldsymbol{b}_1 + \boldsymbol{b}_2\\
&= (0,0,0,2,2,2,1,1,1) + (0,1,2,0,1,2,0,1,2)\\
&= (0,1,2,2,0,1,1,2,0).
\end{aligned}$$

We can represent this resolution 3 fractional factorial design as a matrix F, shown in Table 6.1. We will use each row of F to correspond to an item that will be presented to the respondent in the first position of a choice set. □

■ **EXAMPLE 6.1.2.**

In this example, we construct a 2^{7-4} starting design in 8 runs. We construct this design starting with the 2^3 complete factorial design, and adjoin four attributes which we define as linear

F			
0	0	0	0
0	1	1	1
0	2	2	2
1	0	1	2
1	1	2	0
1	2	0	1
2	0	2	1
2	1	0	2
2	2	1	0

Table 6.1: The 3^{4-2} regular fractional factorial design constructed in Example 6.1.1

combinations of the first three attributes.

$$\begin{aligned}
\boldsymbol{b}_4 &= 0 \times \boldsymbol{b}_1 + 1 \times \boldsymbol{b}_2 + 1 \times \boldsymbol{b}_3 \\
&= (0,0,1,1,0,0,1,1) + (0,1,0,1,0,1,0,1) \\
&= (0,1,1,0,0,1,1,0) \\
\boldsymbol{b}_5 &= 1 \times \boldsymbol{b}_1 + 0 \times \boldsymbol{b}_2 + 1 \times \boldsymbol{b}_3 \\
&= (0,0,0,0,1,1,1,1) + (0,1,0,1,0,1,0,1) \\
&= (0,1,0,1,1,0,1,0) \\
\boldsymbol{b}_6 &= 1 \times \boldsymbol{b}_1 + 1 \times \boldsymbol{b}_2 + 1 \times \boldsymbol{b}_3 \\
&= (0,0,0,0,1,1,1,1) + (0,0,1,1,0,0,1,1) + (0,1,0,1,0,1,0,1) \\
&= (0,1,1,0,1,0,0,1) \\
\boldsymbol{b}_7 &= 1 \times \boldsymbol{b}_1 + 1 \times \boldsymbol{b}_2 + 0 \times \boldsymbol{b}_3 \\
&= (0,0,0,0,1,1,1,1) + (0,0,1,1,0,0,1,1) \\
&= (0,0,1,1,1,1,0,0)
\end{aligned}$$

The first three of these additional attributes have the constraint on the coefficient of $\boldsymbol{b}_3$, and the final attribute has 0 as the coefficient of $\boldsymbol{b}_3$ and thus the constraint is placed on the coefficient of $\boldsymbol{b}_2$. All of the additions will be carried out component–wise using the addition rules of $GF[2]$. This gives the design in Table 6.2. □

We can generalise the construction method presented in Examples 6.1.1 and 6.1.2. The defining equations for a fractional factorial design using the Rao–Hamming construction, as

F						
0	0	0	0	0	0	0
0	0	1	1	1	1	0
0	1	0	1	0	1	1
0	1	1	0	1	0	1
1	0	0	0	1	1	1
1	0	1	1	0	0	1
1	1	0	1	1	0	0
1	1	1	0	0	1	0

Table 6.2: The 2^{7-4} starting design constructed in Example 6.1.2

introduced in Section 1.B.3, will be

$$\begin{aligned}
\boldsymbol{b}_{k-p+1} &= \boldsymbol{b}_{k-p-1} + \boldsymbol{b}_{k-p} \\
&\vdots \\
\boldsymbol{b}_{k-p+\ell-1} &= \alpha^{(\ell-1)} \times \boldsymbol{b}_{k-p-1} + \boldsymbol{b}_{k-p} \\
\boldsymbol{b}_{k-p+\ell} &= \boldsymbol{b}_{k-p-2} + \boldsymbol{b}_{k-p} \\
\boldsymbol{b}_{k-p+\ell+1} &= \boldsymbol{b}_{k-p-2} + \boldsymbol{b}_{k-p-1} + \boldsymbol{b}_{k-p} \\
&\vdots \\
\boldsymbol{b}_{k-p+(\ell-1)^{k-1}} &= \alpha^{(\ell-1)} \times \boldsymbol{b}_1 + \ldots + \alpha^{(\ell-1)} \times \boldsymbol{b}_{k-p-1} + \boldsymbol{b}_{k-p} \\
\boldsymbol{b}_{k-p+(\ell-1)^{k-1}+1} &= \boldsymbol{b}_{k-p-2} + \boldsymbol{b}_{k-p-1} \\
&\vdots \\
\boldsymbol{b}_{k-p+(\ell-1)^{k-1}+(\ell-1)^{k-2}} &= \alpha^{(\ell-1)} \times \boldsymbol{b}_1 + \ldots + \alpha^{(\ell-1)} \times \boldsymbol{b}_{k-p-2} + \boldsymbol{b}_{k-p-1} \\
&\vdots \\
\boldsymbol{b}_k &= \alpha^{(\ell-1)} \times \boldsymbol{b}_1 + \boldsymbol{b}_2,
\end{aligned}$$

with addition and multiplication performed component–wise using the rules of $GF[\ell]$, whose root is denoted by α. An arbitrary entry in the vector $\boldsymbol{b}_q$, $b_{q,i}$, which corresponds to the level of the q^{th} attribute in the i^{th} item is given by

$$b_{q,i} = \sum_{j=1}^{k-p} a_{q,j} \times b_{j,i},$$

subject to the constraint that the last non–zero $a_{q,j}$ be equal to 1. As was the case in the examples, let F be the set of items generated from this construction.

We can describe $\boldsymbol{b}_{k-p+1}, \ldots, \boldsymbol{b}_k$ in terms of permutation matrices. Suppose that for the q^{th} attribute, attribute h, $2 \le h \le k-p$, has the last non–zero coefficient in the defining equation for $\boldsymbol{b}_q$. Then we begin the construction of $\boldsymbol{b}_q$ with $\boldsymbol{b}_h$, which has ℓ^{h-1} repetitions of

$$(0, 1, 2, \ldots, \ell-1) \otimes \boldsymbol{j}_\ell^T \otimes \ldots \otimes \boldsymbol{j}_\ell^T, \qquad (6.1)$$

where there are $(k-p-h-1)$ $\boldsymbol{j}_\ell^T$s. We then add

$$\sum_{j=1}^{h-1} a_{q,j} b_{j,i}$$

component–wise in $GF[\ell]$. Within each repetition of the expression in Equation 6.1, the value of this sum will be the same, since there are more $\boldsymbol{j}_\ell^T$s following the $(0,1,\ldots \ell-1)$ in the expressions for $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_{h-1}$ than in the expression for $\boldsymbol{b}_h$. Therefore we can post–multiply $(0,1,\ldots,\ell-1)$ in each repetition of Equation 6.1 by a permutation matrix that reflects the addition of $\sum_{j=1}^{h-1} a_{q,j} b_{j,i}$.

Since these additions are being carried out in $GF[\ell]$, we define Q_i to be an $\ell \times \ell$ permutation matrix where

$$(Q_i)_{xy} = \begin{cases} 1, & \text{if } x+i=y \text{ in } GF[\ell], \\ 0, & \text{otherwise.} \end{cases}$$

We note that Q_i is a permutation matrix that reflects addition in $GF[\ell]$, whereas P_i, as introduced in Section 1.2, is a permutation matrix that reflects addition modulo ℓ. These matrices are the same only if ℓ is prime.

In the next two examples, we illustrate the construction of a fractional factorial design using the permutation matrices Q_i.

■ **EXAMPLE 6.1.3.**

Consider the 3^{4-2} choice experiment introduced in Example 6.1.1. In this example we constructed a resolution 3 fractional factorial design using

$$\begin{aligned} \boldsymbol{b}_1 &= (0,0,0,1,1,1,2,2,2), \\ \boldsymbol{b}_2 &= (0,1,2,0,1,2,0,1,2), \\ \boldsymbol{b}_3 &= (0,1,2,1,2,0,2,0,1), \text{ and} \\ \boldsymbol{b}_4 &= (0,1,2,2,0,1,1,2,0). \end{aligned}$$

Notice that the third and fourth attributes do not contain three repetitions of $(0,1,2)$, but instead three permutations of $(0,1,2)$. We can use 3×3 permutation matrices to permute the columns of $(0,1,2)$ to form $\boldsymbol{b}_3$ and $\boldsymbol{b}_4$, giving

$$\begin{aligned} \boldsymbol{b}_3 &= \Big((0,1,2) \cdot Q_{1\times b_{1,1}}, (0,1,2) \cdot Q_{1\times b_{1,2}}, (0,1,2) \cdot Q_{1\times b_{1,3}}\Big) \\ &= \Big((0,1,2) \cdot Q_{1\times 0}, (0,1,2) \cdot Q_{1\times 1}, (0,1,2) \cdot Q_{1\times 2}\Big) \\ &= \Big((0,1,2) \cdot Q_0, (0,1,2) \cdot Q_1, (0,1,2) \cdot Q_2\Big) \\ &= ((0,1,2),(1,2,0),(2,0,1)) \\ \boldsymbol{b}_4 &= \Big((0,1,2) \cdot Q_{2\times b_{1,1}}, (0,1,2) \cdot Q_{2\times b_{1,2}}, (0,1,2) \cdot Q_{2\times b_{1,3}}\Big) \\ &= \Big((0,1,2) \cdot Q_{2\times 0}, (0,1,2) \cdot Q_{2\times 1}, (0,1,2) \cdot Q_{2\times 2}\Big) \\ &= \Big((0,1,2) \cdot Q_0, (0,1,2) \cdot Q_2, (0,1,2) \cdot Q_1\Big) \\ &= ((0,1,2),(2,0,1),(1,2,0)) \end{aligned}$$

□

■ **EXAMPLE 6.1.4.**

In this example, we use permutation matrices to construct the last four attributes of the 2^{7-4} fractional factorial design in Example 6.1.2 from the first three attributes.

For the fourth, fifth and sixth attributes, we can see that the corresponding columns of Table 6.2 contain four copies of $(0,1)$, each permuted based on the values of the first two attributes, and their contribution in the defining equations for these attributes. We can define 2×2 permutation matrices to permute each $(0,1)$ to obtain

$$\begin{aligned}
\boldsymbol{b}_4 &= \Big((0,1)\cdot Q_{0\times b_{1,1}+1\times b_{2,1}}, (0,1)\cdot Q_{0\times b_{1,2}+1\times b_{2,2}}, (0,1)\cdot Q_{0\times b_{1,3}+1\times b_{2,3}}, (0,1)\cdot Q_{0\times b_{1,4}+1\times b_{2,4}}\Big)\\
&= \Big((0,1)\cdot Q_{0\times 0+1\times 0}, (0,1)\cdot Q_{0\times 0+1\times 1}, (0,1)\cdot Q_{0\times 1+1\times 0}, (0,1)\cdot Q_{0\times 1+1\times 1}\Big)\\
&= \Big((0,1)\cdot Q_0, (0,1)\cdot Q_1, (0,1)\cdot Q_0, (0,1)\cdot Q_1\Big)\\
&= ((0,1),(1,0),(0,1),(1,0)),\\
\boldsymbol{b}_5 &= \Big((0,1)\cdot Q_{1\times b_{1,1}+0\times b_{2,1}}, (0,1)\cdot Q_{1\times b_{1,2}+0\times b_{2,2}}, (0,1)\cdot Q_{1\times b_{1,3}+0\times b_{2,3}}, (0,1)\cdot Q_{1\times b_{1,4}+0\times b_{2,4}}\Big)\\
&= \Big((0,1)\cdot Q_{1\times 0+0\times 0}, (0,1)\cdot Q_{1\times 0+0\times 1}, (0,1)\cdot Q_{1\times 1+0\times 0}, (0,1)\cdot Q_{1\times 1+0\times 1}\Big)\\
&= \Big((0,1)\cdot Q_0, (0,1)\cdot Q_0, (0,1)\cdot Q_1, (0,1)\cdot Q_1\Big)\\
&= ((0,1),(0,1),(1,0),(1,0)),
\end{aligned}$$

and

$$\begin{aligned}
\boldsymbol{b}_6 &= \Big((0,1)\cdot Q_{1\times b_{1,1}+1\times b_{2,1}}, (0,1)\cdot Q_{1\times b_{1,2}+1\times b_{2,2}}, (0,1)\cdot Q_{1\times b_{1,3}+1\times b_{2,3}}, (0,1)\cdot Q_{1\times b_{1,4}+1\times b_{2,4}}\Big)\\
&= \Big((0,1)\cdot Q_{1\times 0+1\times 0}, (0,1)\cdot Q_{1\times 0+1\times 1}, (0,1)\cdot Q_{1\times 1+1\times 0}, (0,1)\cdot Q_{1\times 1+1\times 1}\Big)\\
&= \Big((0,1)\times Q_0, (0,1)\times Q_1, (0,1)\times Q_1, (0,1)\times Q_0\Big)\\
&= ((0,1),(1,0),(1,0),(0,1)),
\end{aligned}$$

where the first entry in each permutation matrix accounts for the effect of the addition of a multiple of $\boldsymbol{b}_1$ to $(0,1)$ and the second entry in each permutation matrix accounts for the addition of a multiple of $\boldsymbol{b}_2$ to $(0,1)$.

For the final attribute, we notice that the coefficient of $\boldsymbol{b}_3$ in the defining equation is 0, and the attribute with the last non–zero coefficient is the second attribute, that is $h=2$. We can then break the entries for this attribute into

$$\begin{aligned}
\boldsymbol{b}_7 &= \Big(((0,1)\times Q_{1\times b_{1,1}})\otimes \boldsymbol{j}_2^T, ((0,1)\times Q_{1\times b_{1,2}})\otimes \boldsymbol{j}_2^T\Big)\\
&= \Big(((0,1)\times Q_{1\times 0})\otimes \boldsymbol{j}_2^T, ((0,1)\times Q_{1\times 1})\otimes \boldsymbol{j}_2^T\Big)\\
&= \Big(((0,1)\times Q_0)\otimes \boldsymbol{j}_2^T, ((0,1)\times Q_1)\otimes \boldsymbol{j}_2^T\Big)\\
&= \Big((0,1)\otimes \boldsymbol{j}_2^T, (1,0)\otimes \boldsymbol{j}_2^T\Big)\\
&= ((0,0,1,1),(1,1,0,0)),
\end{aligned}$$

where the permutation matrices account for the addition of a multiple of $\boldsymbol{b}_1$. □

This procedure also works if ℓ is a prime power as the next example illustrates.

■ **EXAMPLE 6.1.5.**

We now use permutation matrices to construct a 4^{5-3} starting design. Since 4 is a prime power, $GF[4]$ exists with elements 0, 1, α, and $\alpha^2=\alpha+1$. We can represent addition in $GF[4]$ in terms of the permutation matrices

$$Q_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad Q_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$Q_\alpha = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \qquad Q_{\alpha^2} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Notice that these permutation matrices are different from permutation matrices that reflect addition modulo 4, which would be cyclic. These permutation matrices are in fact the Kronecker product of two 2×2 cyclic permutation matrices.

Using the modified Rao–Hamming construction, we have the following defining equations, which we will use to construct the starting design.

$$\begin{aligned} \boldsymbol{b}_3 &= 1 \times \boldsymbol{b}_1 + 1 \times \boldsymbol{b}_2, \\ \boldsymbol{b}_4 &= \alpha \times \boldsymbol{b}_1 + 1 \times \boldsymbol{b}_2, \end{aligned}$$

and

$$\boldsymbol{b}_5 = \alpha^2 \times \boldsymbol{b}_1 + 1 \times \boldsymbol{b}_2,$$

where

$$\begin{aligned} \boldsymbol{b}_1 &= (0,1,\alpha,\alpha^2) \otimes \boldsymbol{j}_4^T \\ &= (0,0,0,0,1,1,1,1,\alpha,\alpha,\alpha,\alpha,\alpha^2,\alpha^2,\alpha^2,\alpha^2), \end{aligned}$$

and

$$\begin{aligned} \boldsymbol{b}_2 &= \boldsymbol{j}_4^T \otimes (0,1,\alpha,\alpha^2) \\ &= (0,1,\alpha,\alpha^2,0,1,\alpha,\alpha^2,0,1,\alpha,\alpha^2,0,1,\alpha,\alpha^2). \end{aligned}$$

Then, in the same way as Examples 6.1.3 and 6.1.4, we use the defining equations to permute the entries of $\boldsymbol{b}_2$ to obtain the remaining attributes. So we have

$$\begin{aligned} \boldsymbol{b}_3 &= \Big((0,1,\alpha,\alpha^2)\cdot Q_{1\times b_{1,1}}, (0,1,\alpha,\alpha^2)\cdot Q_{1\times b_{1,2}}, (0,1,\alpha,\alpha^2)\cdot Q_{1\times b_{1,3}}, (0,1,\alpha,\alpha^2)\cdot Q_{1\times b_{1,4}}\Big) \\ &= \Big((0,1,\alpha,\alpha^2)\cdot Q_{1\times 0}, (0,1,\alpha,\alpha^2)\cdot Q_{1\times 1}, (0,1,\alpha,\alpha^2)\cdot Q_{1\times \alpha}, (0,1,\alpha,\alpha^2)\cdot Q_{1\times \alpha^2}\Big) \\ &= \Big((0,1,\alpha,\alpha^2)\cdot Q_0, (0,1,\alpha,\alpha^2)\cdot Q_1, (0,1,\alpha,\alpha^2)\cdot Q_\alpha, (0,1,\alpha,\alpha^2)\cdot Q_{\alpha^2}\Big) \\ &= ((0,1,\alpha,\alpha^2),(1,0,\alpha^2,\alpha),(\alpha,\alpha^2,0,1),(\alpha^2,\alpha,1,0)), \\ \boldsymbol{b}_4 &= \Big((0,1,\alpha,\alpha^2)\cdot Q_{\alpha\times b_{1,1}}, (0,1,\alpha,\alpha^2)\cdot Q_{\alpha\times b_{1,2}}, (0,1,\alpha,\alpha^2)\cdot Q_{\alpha\times b_{1,3}}, (0,1,\alpha,\alpha^2)\cdot Q_{\alpha\times b_{1,4}}\Big) \\ &= \Big((0,1,\alpha,\alpha^2)\cdot Q_{\alpha\times 0}, (0,1,\alpha,\alpha^2)\cdot Q_{\alpha\times 1}, (0,1,\alpha,\alpha^2)\cdot Q_{\alpha\times \alpha}, (0,1,\alpha,\alpha^2)\cdot Q_{\alpha\times \alpha^2}\Big) \\ &= \Big((0,1,\alpha,\alpha^2)\cdot Q_0, (0,1,\alpha,\alpha^2)\cdot Q_\alpha, (0,1,\alpha,\alpha^2)\cdot Q_{\alpha^2}, (0,1,\alpha,\alpha^2)\cdot Q_1\Big) \\ &= ((0,1,\alpha,\alpha^2),(\alpha,\alpha^2,0,1),(\alpha^2,\alpha,1,0),(1,0,\alpha^2,\alpha)), \end{aligned}$$

and

$$\begin{aligned}
\boldsymbol{b}_5 &= \Big((0,1,\alpha,\alpha^2)\cdot Q_{\alpha^2\times b_{1,1}}, (0,1,\alpha,\alpha^2)\cdot Q_{\alpha^2\times b_{1,2}}, (0,1,\alpha,\alpha^2)\cdot Q_{\alpha^2\times b_{1,3}}, (0,1,\alpha,\alpha^2)\cdot Q_{\alpha^2\times b_{1,4}}\Big)\\
&= \Big((0,1,\alpha,\alpha^2)\cdot Q_{\alpha^2\times 0}, (0,1,\alpha,\alpha^2)\cdot Q_{\alpha^2\times 1}, (0,1,\alpha,\alpha^2)\cdot Q_{\alpha^2\times \alpha}, (0,1,\alpha,\alpha^2)\cdot Q_{\alpha^2\times \alpha^2}\Big)\\
&= \Big((0,1,\alpha,\alpha^2)\cdot Q_0, (0,1,\alpha,\alpha^2)\cdot Q_{\alpha^2}, (0,1,\alpha,\alpha^2)\cdot Q_1, (0,1,\alpha,\alpha^2)\cdot Q_\alpha\Big)\\
&= ((0,1,\alpha,\alpha^2),(\alpha^2,\alpha,1,0),(1,0,\alpha^2,\alpha),(\alpha,\alpha^2,0,1)),
\end{aligned}$$

where the permutation matrices account for the addition of $\boldsymbol{b}_1$. This gives the fractional factorial design in Table 6.3. □

F									
0	0	0	0	0	α	0	α	α^2	1
0	1	1	1	1	α	1	α^2	α	0
0	α	α	α	α	α	α	0	1	α^2
0	α^2	α^2	α^2	α^2	α	α^2	1	0	α
1	0	1	α	α^2	α^2	0	α^2	1	α
1	1	0	α^2	α	α^2	1	α	0	α^2
1	α	α^2	0	1	α^2	α	1	α^2	0
1	α^2	α	1	0	α^2	α^2	0	α	1

Table 6.3: The 4^{5-3} fractional factorial design constructed in Example 6.1.5

Recall that in Section 1.2 we used these vectors of attribute levels as the column labels for contrast matrices. In particular, we repeated the operations that we performed on the vectors of attribute levels on B_F to obtain $B_{F+\boldsymbol{g}_i}$. The next two examples show how this idea can be used to define the rows of a contrast matrix for new attributes constructed using the modified Rao–Hamming construction.

■ **EXAMPLE 6.1.6.**

In this example, we use the permutation matrices introduced in Example 6.1.3 to rearrange the columns in the contrast matrix corresponding to the second attribute to obtain the rows of the contrast matrix corresponding to the third and fourth attributes of the fractional factorial design.

Recall from Example 1.B.5 that the normalised contrast matrix corresponding to the main effects of the first two attributes is given by

$$\begin{aligned}
B_{F,(1,2)} &= \begin{bmatrix} B_3 \otimes \frac{1}{\sqrt{3}}\boldsymbol{j}_3^T \\ \frac{1}{\sqrt{3}}\boldsymbol{j}_3^T \otimes B_3 \end{bmatrix}\\
&= \frac{1}{\sqrt{3}} \begin{array}{c} \begin{matrix} \scriptstyle 00 & \scriptstyle 01 & \scriptstyle 02 & \scriptstyle 10 & \scriptstyle 11 & \scriptstyle 12 & \scriptstyle 20 & \scriptstyle 21 & \scriptstyle 22 \end{matrix} \\ \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} \end{array}.
\end{aligned}$$

The columns of the contrast matrix can be labelled by the items in the fraction. In particular, if we consider the rows of the contrast matrix corresponding to the main effect of the q^{th} attribute, then we can labels the columns by the level that that attribute takes in the corresponding item. So we can permute the rows of the contrast matrix corresponding to the second attribute to obtain the rows of the contrast matrix corresponding to the third and fourth attributes in the same way as we permuted the levels vectors in Example 6.1.3. The rows of the contrast matrix corresponding to the last attribute in $\boldsymbol{b}_q$ with a non–zero coefficient, which in this case the second attribute, can be expressed as

$$B_{F(2)} = \frac{1}{\sqrt{3}}\left[\begin{array}{ccc} B_3 & B_3 & B_3 \end{array}\right],$$

with each B_3 corresponding to the contrast coefficients for the levels of one replication of $(0,1,2)$. Then, in the same way that we used permutation matrices to change the order of $(0,1,2)$, we can post–multiply each B_3 by a permutation matrix to permute the columns of B_3 to be consistent with the entries in $\boldsymbol{b}_3$ and $\boldsymbol{b}_4$, giving the contrast matrix $B_{F,(3,4)}$.

$$\begin{aligned}
B_{F,(3,4)} &= \frac{1}{\sqrt{3}}\left[\begin{array}{ccc} B_3Q_{1\times b_{1,1}} & B_3Q_{1\times b_{1,2}} & B_3Q_{1\times b_{1,3}} \\ B_3Q_{2\times b_{1,1}} & B_3Q_{2\times b_{1,2}} & B_3Q_{2\times b_{1,3}} \end{array}\right] \\
&= \frac{1}{\sqrt{3}}\left[\begin{array}{ccc} B_3Q_0 & B_3Q_1 & B_3Q_2 \\ B_3Q_0 & B_3Q_2 & B_3Q_1 \end{array}\right] \\
&= \frac{1}{\sqrt{3}}\left[\begin{array}{ccccccccc}
{\scriptstyle 00} & {\scriptstyle 11} & {\scriptstyle 22} & {\scriptstyle 12} & {\scriptstyle 20} & {\scriptstyle 01} & {\scriptstyle 21} & {\scriptstyle 02} & {\scriptstyle 10} \\
\frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\
\frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} \\
\frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\
\frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}}
\end{array}\right].
\end{aligned}$$

The full contrast matrix for B_F will be a matrix containing the rows of the contrast matrix corresponding to each of the attributes, giving

$$B_F = \frac{1}{\sqrt{3}}\left[\begin{array}{ccccccccc}
{\scriptstyle 0000} & {\scriptstyle 0111} & {\scriptstyle 0222} & {\scriptstyle 1012} & {\scriptstyle 1120} & {\scriptstyle 1201} & {\scriptstyle 2021} & {\scriptstyle 2102} & {\scriptstyle 2210} \\
\frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\
\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\
\frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\
\frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\
\frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\
\frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} \\
\frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\
\frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}}
\end{array}\right].$$

□

In the next example we use the same procedure to construct B_F for the 2^{7-4} design introduced in Example 6.1.2.

■ **EXAMPLE 6.1.7.**

Consider the 2^{7-4} fractional factorial design introduced in Example 6.1.2. The contrast matrix for the main effects of the first three attributes will form the contrast matrix $B_{F,1}$ is

$$
\begin{aligned}
B_{F,(1,2,3)} &= \begin{bmatrix} B_2 \otimes \frac{1}{\sqrt{2}}\boldsymbol{j}_2^T \otimes \frac{1}{\sqrt{2}}\boldsymbol{j}_2^T \\ \frac{1}{\sqrt{2}}\boldsymbol{j}_2^T \otimes B_2 \otimes \frac{1}{\sqrt{2}}\boldsymbol{j}_2^T \\ \frac{1}{\sqrt{2}}\boldsymbol{j}_2^T \otimes \frac{1}{\sqrt{2}}\boldsymbol{j}_2^T \otimes B_2 \end{bmatrix} \\
&= \begin{array}{c} \begin{array}{cccccccc} 000 & 001 & 010 & 011 & 100 & 101 & 110 & 111 \end{array} \\ \begin{bmatrix} \frac{-1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} \\ \frac{-1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} \\ \frac{-1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} \end{bmatrix} \end{array} .
\end{aligned}
$$

We can use the relationships developed in Example 6.1.4 to adapt the row of the contrast matrix corresponding to the third attribute to obtain the rows corresponding to the fourth, fifth and sixth attributes. First we note that the row of the contrast matrix corresponding to the third attribute can be expressed as

$$B_{F,(3)} = \frac{1}{2}\begin{bmatrix} B_2 & B_2 & B_2 & B_2 \end{bmatrix},$$

where each B_2 corresponding to the contrast coefficients for one replication of $(0,1)$ in $\boldsymbol{b}_3$. Then we can permute the columns of B_2 in the same way as the $(0,1)$s, giving the rows of the contrast matrix corresponding to the fourth, fifth and sixth attributes B_F;

$$
\begin{aligned}
B_{F,(4,5,6)} &= \frac{1}{2}\begin{bmatrix} B_2Q_{0\times b_{1,1}+1\times b_{1,2}} & B_2Q_{0\times b_{2,1}+1\times b_{2,2}} & B_2Q_{0\times b_{1,3}+1\times b_{2,3}} & B_2Q_{0\times b_{1,4}+1\times b_{2,4}} \\ B_2Q_{1\times b_{1,1}+0\times b_{1,2}} & B_2Q_{1\times b_{2,1}+0\times b_{2,2}} & B_2Q_{1\times b_{1,3}+0\times b_{2,3}} & B_2Q_{1\times b_{1,4}+0\times b_{2,4}} \\ B_2Q_{1\times b_{1,1}+1\times b_{1,2}} & B_2Q_{1\times b_{2,1}+1\times b_{2,2}} & B_2Q_{1\times b_{1,3}+1\times b_{2,3}} & B_2Q_{1\times b_{1,4}+1\times b_{2,4}} \end{bmatrix} \\
&= \frac{1}{2}\begin{bmatrix} B_2Q_{0\times 0+1\times 0} & B_2Q_{0\times 0+1\times 1} & B_2Q_{0\times 1+1\times 0} & B_2Q_{0\times 1+1\times 1} \\ B_2Q_{1\times 0+0\times 0} & B_2Q_{1\times 0+0\times 1} & B_2Q_{1\times 1+0\times 0} & B_2Q_{1\times 1+0\times 1} \\ B_2Q_{1\times 0+1\times 0} & B_2Q_{1\times 0+1\times 1} & B_2Q_{1\times 1+1\times 0} & B_2Q_{1\times 1+1\times 1} \end{bmatrix} \\
&= \frac{1}{2}\begin{bmatrix} B_2Q_0 & B_2Q_1 & B_2Q_0 & B_2Q_1 \\ B_2Q_0 & B_2Q_0 & B_2Q_1 & B_2Q_1 \\ B_2Q_0 & B_2Q_1 & B_2Q_1 & B_2Q_0 \end{bmatrix} \\
&= \begin{array}{c} \begin{array}{cccccccc} 000 & 111 & 101 & 010 & 011 & 100 & 110 & 001 \end{array} \\ \begin{bmatrix} \frac{-1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} \\ \frac{-1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} \\ \frac{-1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1-}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} \end{bmatrix} \end{array} .
\end{aligned}
$$

We also notice that the row of the contrast matrix corresponding to the second attribute can be expressed as

$$B_{F,(2)} = \frac{1}{2}\begin{bmatrix} B_2 \otimes \boldsymbol{j}_2^T & B_2 \otimes \boldsymbol{j}_2^T \end{bmatrix}.$$

Then we can permute the columns of B_2 in the same way as the $((0,1)\otimes \boldsymbol{j}_2^T)$s in Example 6.1.4, giving the rows of the contrast matrix corresponding to the final attribute $B_{F,(7)}$.

$$\begin{aligned}
B_{F,(7)} &= \frac{1}{\sqrt{2}}\left[\begin{array}{cc}(B_2Q_{1\times b_{1,1}})\otimes\frac{1}{\sqrt{2}}\boldsymbol{j}_2^T & (B_2Q_{1\times b_{1,2}})\otimes\frac{1}{\sqrt{2}}\boldsymbol{j}_2^T\end{array}\right]\\
&= \frac{1}{\sqrt{2}}\left[\begin{array}{cc}(B_2Q_{1\times 0})\otimes\frac{1}{\sqrt{2}}\boldsymbol{j}_2^T & (B_2Q_{1\times 1})\otimes\frac{1}{\sqrt{2}}\boldsymbol{j}_2^T\end{array}\right]\\
&= \frac{1}{2}\left[\begin{array}{cc}(\frac{-1}{\sqrt{2}},\frac{1}{\sqrt{2}})\otimes(1,1) & (\frac{1}{\sqrt{2}},\frac{-1}{\sqrt{2}})\otimes(1,1)\end{array}\right]\\
&= \begin{array}{c}\begin{array}{cccccccc}0 & 0 & 1 & 1 & 1 & 1 & 0 & 0\end{array}\\ \left[\begin{array}{cccccccc}\frac{-1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}} & \frac{-1}{2\sqrt{2}}\end{array}\right]\end{array}.
\end{aligned}$$

Putting all of this together, we obtain the contrast matrix for this fractional factorial design. □

The constructions in these two examples can be generalised for any ℓ^{k-p} regular fractional factorial design where ℓ is a prime or a prime power. For each ℓ level attribute there will be $\ell-1$ associated contrasts to describe the components of the main effect. Then the contrast for a particular attribute can be represented as a row of block matrices. We construct this row of block matrices based on the defining equation for the attribute. The form of this row will depend on the value of h, where the h^{th} attribute is the last attribute in the defining equation with a non–zero coefficient.

For the first $(\ell-1)^{k-p-1}$ attributes, $\boldsymbol{b}_{k-p}$ will have the last non–zero coefficient in the defining equation. Thus the row of block matrices of the contrast matrix corresponding to one of these $(\ell-1)^{k-p-1}$ attributes is given by

$$\begin{aligned}
&\frac{1}{\ell^{(k-p-1)/2}}\left[B_\ell Q_{(0\times a_{q,1}+\ldots+0\times a_{q,k-p-1})}, B_\ell Q_{(0\times a_{q,1}+\ldots+0\times a_{q,k-p-2}+1\times a_{q,k-p-1})},\ldots,\right.\\
&\qquad\left. B_\ell Q_{(0\times a_{q,1}+\ldots+0\times a_{q,k-p-2}+\alpha^{\ell-2}\times a_{q,k-p-1})},\ldots, B_\ell Q_{(\alpha^{\ell-2}\times a_{q,1}+\ldots+\alpha^{\ell-2}\times a_{q,k-p-1})}\right].
\end{aligned}$$

For the next $(\ell-1)^{k-p-2}$ attributes, the last non–zero coefficient in the defining equation will be the coefficient of $\boldsymbol{b}_{k-p-1}$. Then the row of block matrices of the contrast matrix corresponding to one of these $(\ell-1)^{k-p-2}$ attributes is given by

$$\begin{aligned}
\frac{1}{\ell^{(k-p-2)/2}}&\left[(B_\ell Q_{(0\times a_{q,1}+\ldots+0\times a_{q,k-p-2})})\otimes\frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T,\right.\\
&(B_\ell Q_{(0\times a_{q,1}+\ldots+0\times a_{q,k-p-3}+1\times a_{q,k-p-2})})\otimes\frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T,\ldots,\\
&(B_\ell Q_{(0\times a_{q,1}+\ldots+0\times a_{q,k-p-3}+\alpha^{\ell-2}\times a_{q,k-p-2})})\otimes\frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T,\ldots,\\
&\left.(B_\ell Q_{(\alpha^{\ell-2}\times a_{q,1}+\ldots+\alpha^{\ell-2}\times a_{q,k-p-2})})\otimes\frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T\right].
\end{aligned}$$

This construction can be continued until the final $\ell-1$ attributes, whose last non–zero coefficient in their defining equation will be $\boldsymbol{b}_2$. Then the row of block matrices of the contrast matrix corresponding to one of these $\ell-1$ attributes is given by

$$\begin{aligned}
\frac{1}{\ell^{1/2}}&\left[(B_\ell Q_{(0\times a_{q,1})})\otimes\frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T\otimes\ldots\otimes\frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T,(B_\ell Q_{(1\times a_{q,1})})\otimes\frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T\otimes\ldots\otimes\frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T,\right.\\
&\left.(B_\ell Q_{(\alpha\times a_{q,1})})\otimes\frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T\otimes\ldots\otimes\frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T,\ldots,(B_\ell Q_{(\alpha^{\ell-2}\times a_{q,1})})\otimes\frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T\otimes\ldots\otimes\frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T\right].
\end{aligned}$$

Putting the complete and fractional components together, the contrast matrix for the first fractional factorial design which gives rise to the options in each of the choice sets in the experiment can be expressed as

$$B_F = \begin{bmatrix} B_{F,(1)} \\ B_{F,(2)} \\ \vdots \\ B_{F,(k)} \end{bmatrix}.$$

Now that we have constructed a contrast matrix for the items that appear in the first position of each choice set, we use this to construct the contrast matrices for the items that appear in the remaining positions of the choice set.

6.2 Adding Generators to a Starting Design

In this section, we consider how to transform the contrast matrix of a fractional factorial design to incorporate the consequences of changing the items in the fraction caused by the addition of a generator to each row of that design. The method that will be used to do this is the same as the method we used in Section 1.2 when a full factorial starting design was used. That is, we can define a permutation matrix $Q_{g_{i,q}}$ that will, when B_F is post–multiplied by the permutation matrix, permute the columns of B_F in such a way that it will reflect the change in items caused by the addition of a generator $\boldsymbol{g}_i = [g_{i,q}]$ in $GF[\ell]$ to the starting design.

Suppose that $m-1$ generators are added to the rows of F to obtain m fractional factorial designs. We form the choice sets by presenting items described by the same row of each of the m fractional factorial designs. There are $N = \ell^{k-p}$ such choice sets. Let $G = (\boldsymbol{g}_1 = \boldsymbol{0}, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_m)$, and for each of these $\boldsymbol{g}_i$, let $\boldsymbol{g}_i = (g_{i,1}, g_{i,2}, \ldots, g_{i,k})$. Then the choice sets become the rows of

$$[F, F+\boldsymbol{g}_2, F+\boldsymbol{g}_3, \ldots, F+\boldsymbol{g}_m].$$

We now consider a small example of how we can modify the contrast matrix to incorporate the addition of a generator. We will use the permutation matrices introduced in Section 6.1 to do this.

■ **EXAMPLE 6.2.1.**

Suppose that we begin with the design developed in Example 6.1.1. In Example 6.1.6 we found the contrast matrix for the items that appear in the first positions of each choice set. Now suppose that we add the generator $\boldsymbol{g}_2 = (1212)$ to obtain a second design, as shown by the second set of columns of the design in Table 6.4.

To obtain the contrast matrix for the second design, $F+\boldsymbol{g}_2$, we post–multiply each occurrence of B_3 by a permutation matrix to reorder the columns of B_3, as we did in Example 1.2.7. The choice of the permutation matrix will depend on which attribute the row of blocks containing B_3 corresponds to.

B_3 is post-multiplied by the chosen permutation matrix. This will permute the columns of B_3 in a way that reflects the changed order caused by the addition of the generator $\boldsymbol{g}_2$. Then each occurrence of B_3 in the row of blocks corresponding to attribute q will be post–multiplied by the permutation matrix $Q_{g_{2,q}}$. Thus we obtain

$$
\begin{aligned}
B_{F+\boldsymbol{g}_2} &= \begin{bmatrix} (B_3Q_{g_{2,1}}) \otimes \frac{1}{\sqrt{3}}\boldsymbol{j}_3^T \\ \frac{1}{\sqrt{3}}\boldsymbol{j}_3^T \otimes (B_3Q_{g_{2,2}}) \\ \frac{1}{\sqrt{3}}(B_3Q_{g_{2,3}}Q_0 \quad B_3Q_{g_{2,3}}Q_1 \quad B_3Q_{g_{2,3}}Q_2) \\ \frac{1}{\sqrt{3}}(B_3Q_{g_{2,4}}Q_0 \quad B_3Q_{g_{2,4}}Q_2 \quad B_3Q_{g_{2,4}}Q_1) \end{bmatrix} \\
&= \begin{bmatrix} (B_3Q_1) \otimes \frac{1}{\sqrt{3}}\boldsymbol{j}_3^T \\ \frac{1}{\sqrt{3}}\boldsymbol{j}_3^T \otimes (B_3Q_2) \\ \frac{1}{\sqrt{3}}(B_3Q_1Q_0 \quad B_3Q_1Q_1 \quad B_3Q_1Q_2) \\ \frac{1}{\sqrt{3}}(B_3Q_2Q_0 \quad B_3Q_2Q_2 \quad B_3Q_2Q_1) \end{bmatrix} \\
&= \begin{bmatrix} (B_3Q_1) \otimes \frac{1}{\sqrt{3}}\boldsymbol{j}_3^T \\ \frac{1}{\sqrt{3}}\boldsymbol{j}_3^T \otimes (B_3Q_2) \\ \frac{1}{\sqrt{3}}(B_3Q_1 \quad B_3Q_2 \quad B_3Q_0) \\ \frac{1}{\sqrt{3}}(B_3Q_2 \quad B_3Q_1 \quad B_3Q_0) \end{bmatrix}
\end{aligned}
$$

$$
= \frac{1}{\sqrt{3}} \begin{bmatrix}
\scriptstyle 1212 & \scriptstyle 1020 & \scriptstyle 1101 & \scriptstyle 2221 & \scriptstyle 2002 & \scriptstyle 2110 & \scriptstyle 0200 & \scriptstyle 0011 & \scriptstyle 0122 \\
0 & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\
\frac{-2}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\
\frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\
\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} \\
0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\
\frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\
\frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\
\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}}
\end{bmatrix}. \tag{6.2}
$$

□

This method can be generalised to allow any regular fractional factorial design to be used as the starting design. We begin by considering the attributes that were chosen in Section 6.1 to form an ℓ^{k-p} complete factorial when the columns and rows of the starting design were reordered. Let the contrast matrix for these attributes form $B_{F,1}$. If a generator $\boldsymbol{g}_i$ is added to F, then the contrast matrix $B_{F,1}$ becomes

$$
B_{F+\boldsymbol{g}_i,1} = \begin{bmatrix}
B_\ell Q_{g_{i,1}} \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes \ldots \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \\
\frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes B_\ell Q_{g_{i,2}} \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes \ldots \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \\
\vdots \\
\frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes \ldots \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes B_\ell Q_{g_{i,k-p}}
\end{bmatrix}.
$$

We can do the same thing for the remaining p attributes. We do this by post–multiplying each occurrence of the contrast matrix B_ℓ by a permutation matrix $Q_{g_{i,q}}$, permuting the entries

F				$\mathbf{F}+\mathbf{g_2}$			
0	0	0	0	1	2	1	2
0	1	1	1	1	0	2	0
0	2	2	2	1	1	0	1
1	0	1	2	2	2	2	1
1	1	2	0	2	0	0	2
1	2	0	1	2	1	1	0
2	0	2	1	0	2	0	0
2	1	0	2	0	0	1	1
2	2	1	0	0	1	2	2

Table 6.4: Adding the generator (1212) to the design in Example 6.1.1

of the contrast matrix to reflect the effect of adding a generator to the starting design. Again the form of the row of block matrices corresponding to the q^{th} attribute, where $k-p+1 \leq q \leq k$, depends on the value of h, the attribute in the defining equation for attribute q with the last non–zero coefficient.

For the rows of block matrices corresponding to the first $(\ell-1)^{k-p-1}$ of these attributes, the j^{th} block matrix in the row of block matrices corresponding to the q^{th} attribute is given by

$$(B_{F+\boldsymbol{g}_i})_{q,j} = \frac{1}{\ell^{(k-p-1)/2}} B_\ell \times Q_{g_{i,q}} \times Q_{(b_{1,j}a_{q,1}+\ldots+b_{k-p-1,j}\times a_{q,k-p-1})}.$$

For the rows of block matrices corresponding to the next $(\ell-1)^{k-p-2}$ attributes, the j^{th} block matrix in the row of block matrices corresponding to the contrasts of the q^{th} attribute is given by

$$(B_{F+\boldsymbol{g}_i})_{q,j} = \frac{1}{\ell^{(k-p-1)/2}} (B_\ell \times Q_{g_{i,q}} \times Q_{(b_{1,j}a_{q,1}+\ldots+b_{k-p-2,j}a_{q,k-p-2})}) \otimes \boldsymbol{j}_\ell^T.$$

This can be continued until the rows of block matrices corresponding to the final $(\ell-1)$ attributes, the j^{th} block matrix in the row of block matrices corresponding to the contrasts of the q^{th} attribute is given by

$$(B_{F+\boldsymbol{g}_i})_{q,j} = \frac{1}{\ell^{(k-p-1)/2}} (B_\ell \times Q_{g_{i,q}} \times Q_{(b_{1,j}a_{q,1})}) \otimes \boldsymbol{j}_\ell^T \otimes \ldots \otimes \boldsymbol{j}_\ell^T.$$

In order to find an expression for the information matrix for the estimation of main effects, we also need to derive the $\Lambda(\boldsymbol{\pi}_0)$ matrix. The form of the $\Lambda(\boldsymbol{\pi}_0)$ matrix depends on whether the generator $\boldsymbol{g}_i$ forms a row of the starting design F or not. If $\boldsymbol{g}_i \in F$, then $F+\boldsymbol{g}_i \equiv F$, since the addition of a generator to the principal fraction of a regular fractional factorial design is closed under addition in $GF[\ell]$. On the other hand, if $\boldsymbol{g}_i \notin F$, then $F+\boldsymbol{g}_i \neq F$, and the form of the $\Lambda(\boldsymbol{\pi}_0)_{1i}$ matrix will be quite different to the case where $\boldsymbol{g}_i \in F$. We consider these cases in turn.

6.3 Constructing the Information Matrix - Generator does not Appear in the Starting Design

First, we consider the case where $\boldsymbol{g}_i \notin F$. In this case none of the items in F will appear in $F+\boldsymbol{g}_i$, since the principal fraction of a regular fractional factorial design is closed under addition.

Therefore by rearranging the rows of the contrast matrix if necessary the matrix $\Lambda(\boldsymbol{\pi}_0)$, for the pairs of items in positions 1 and i of the choice set, is

$$m^2 N\Lambda(\boldsymbol{\pi}_0)_{1i} = \begin{bmatrix} I_{\ell^{k-p}} & -I_{\ell^{k-p}} & \mathbf{0} \\ -I_{\ell^{k-p}} & I_{\ell^{k-p}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

and the contrast matrix B can be expressed as

$$\frac{1}{\ell^{p/2}} \begin{bmatrix} B_{F,1} & B_{F+\boldsymbol{g}_i,1} & B_{\bar{A},1} \\ B_{F,2} & B_{F+\boldsymbol{g}_i,2} & B_{\bar{A},2} \end{bmatrix},$$

where $B_{\bar{A}}$ contains the contrast matrix for those items that do not appear in either position of the choice set. If $m > 2$ then the argument used in this section can be applied to any pair of positions in the choice set, providing that $\boldsymbol{g}_j - \boldsymbol{g}_i \notin F$.

We will now consider a small example using these expressions to derive the information matrix for a design.

■ **EXAMPLE 6.3.1.**

In this example we have 4 3–level attributes and want to be able to estimate main effects using choice sets of size 3. The contrast matrix for the items in Table 6.1, those that will appear in the first position of each choice set, is given by Equation 6.2. If we add the generators $\boldsymbol{g}_2 = (1111)$ and $\boldsymbol{g}_3 = (2222)$, neither of which appear in the starting design, we obtain the design in Table 6.5. Using the results in Section 6.2, the contrast matrices for $B_{F+\boldsymbol{g}_2}$ and $B_{F+\boldsymbol{g}_3}$ are given by

$$B_{F+\boldsymbol{g}_2} = \begin{bmatrix} & (B_3 \times Q_1) \otimes \frac{1}{\sqrt{3}}\boldsymbol{j}_3 & \\ & \frac{1}{\sqrt{3}}\boldsymbol{j}_3 \otimes (B_3 \times Q_1) & \\ \frac{1}{\sqrt{3}}\Big(B_3 \times Q_1 \times Q_{0\times a_{3,2}} & B_3 \times Q_1 \times Q_{1\times a_{3,2}} & B_3 \times Q_1 \times Q_{2\times a_{3,2}}\Big) \\ \frac{1}{\sqrt{3}}\Big(B_3 \times Q_1 \times Q_{0\times a_{4,2}} & B_3 \times Q_1 \times Q_{1\times a_{4,2}} & B_3 \times Q_1 \times Q_{2\times a_{4,2}}\Big) \end{bmatrix},$$

and

$$B_{F+\boldsymbol{g}_3} = \begin{bmatrix} & (B_3 \times Q_2) \otimes \frac{1}{\sqrt{3}}\boldsymbol{j}_3 & \\ & \frac{1}{\sqrt{3}}\boldsymbol{j}_3 \otimes (B_3 \times Q_2) & \\ \frac{1}{\sqrt{3}}\Big(B_3 \times Q_2 \times Q_{0\times a_{3,2}} & B_3 \times Q_2 \times Q_{1\times a_{3,2}} & B_3 \times Q_2 \times Q_{2\times a_{3,2}}\Big) \\ \frac{1}{\sqrt{3}}\Big(B_3 \times Q_2 \times Q_{0\times a_{4,2}} & B_3 \times Q_2 \times Q_{1\times a_{4,2}} & B_3 \times Q_2 \times Q_{2\times a_{4,2}}\Big) \end{bmatrix}$$

respectively. We can also calculate the $\Lambda(\boldsymbol{\pi}_0)$ matrix for each set of pairs, and thus find the information matrix for the estimation of main effects, $C(\boldsymbol{\pi}_0)_M$, for each pair of options in the same position. For the first two positions of the choice set,

$$\begin{aligned} \Lambda(\boldsymbol{\pi}_0)_{1,2} &= \frac{1}{m^2 N} \begin{bmatrix} I_9 & -I_9 & \mathbf{0} \\ -I_9 & I_9 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \\ &= \frac{1}{81} \begin{bmatrix} I_9 & -I_9 & \mathbf{0} \\ -I_9 & I_9 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}. \end{aligned}$$

Since none of the items are repeated across the choice sets, we can arrange the items in F, $F+\boldsymbol{g}_2$ and $F+\boldsymbol{g}_3$ such that the $\Lambda(\boldsymbol{\pi}_0)$ matrices for each pair of items can be rearranged to have the same form as $\Lambda(\boldsymbol{\pi}_0)_{1,2}$. Then by using the relationship $C(\boldsymbol{\pi}_0)_{i,j} = B\Lambda(\boldsymbol{\pi}_0)_{i,j}B^T$ where B is normalised to give $BB^T = I$, the expression for $C(\boldsymbol{\pi}_0)_{1,2}$ is

$$\begin{aligned} C(\boldsymbol{\pi}_0)_{1,2} &= \frac{1}{3^{2/2}}\left[B_F \quad B_{F+\boldsymbol{g}_2} \quad B_{\bar{A}}\right] \times \frac{1}{81}\begin{bmatrix} I_9 & -I_9 & \mathbf{0} \\ -I_9 & I_9 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \times \frac{1}{3^{2/2}}\begin{bmatrix} B_F^T \\ B_{F+\boldsymbol{g}_2}^T \\ B_{\bar{A}}^T \end{bmatrix} \\ &= \frac{1}{3^6}\left[(B_F - B_{F+\boldsymbol{g}_2}) \times B_F^T + (-B_F + B_{F+\boldsymbol{g}_2}) \times B_{F+\boldsymbol{g}_2}^T\right] \\ &= \frac{1}{3^6} \times (3I_8) \\ &= \frac{1}{3^5} I_8, \end{aligned}$$

with $p = 2$. Similarly, $C(\boldsymbol{\pi}_0)_{1,3} = \frac{1}{3^5}I_8$ and $C(\boldsymbol{\pi}_0)_{2,3} = \frac{1}{3^5}I_8$. Then

$$\begin{aligned} C(\boldsymbol{\pi}_0) &= \sum_{i<j} C(\boldsymbol{\pi}_0)_{i,j} \\ &= \frac{1}{81} I_8. \end{aligned}$$

□

Option 1				**Option 2**				**Option 3**			
0	0	0	0	1	1	1	1	2	2	2	2
1	0	1	1	2	1	2	2	0	2	0	0
2	0	2	2	0	1	0	0	1	2	1	1
0	1	1	2	1	2	2	0	2	1	0	1
1	1	2	0	2	2	0	1	0	1	1	2
2	1	0	1	0	2	1	2	1	1	2	0
0	2	2	1	1	0	0	2	2	0	1	0
1	2	0	2	2	0	1	0	0	0	2	1
2	2	1	0	0	0	2	1	1	0	0	2

Table 6.5: The choice sets used in Example 6.3.2

Instead of multiplying the contrast matrices numerically, we can take advantage of the structure that they have. This will result in an expression which is easier to generalise. We start by returning to Example 6.3.1.

■ **EXAMPLE 6.3.2.**

In this example, we use the structure of the contrast matrices and $\Lambda(\boldsymbol{\pi}_0)$ to derive $C(\boldsymbol{\pi}_0)$ for the design in Table 6.5.

First, notice that the inner product of any two rows of B_F (or $B_{F+\boldsymbol{g}_i}$) that correspond to different attributes will be equal to 0. Thus $B_F B_F^T$ and $B_{F+\boldsymbol{g}_i} B_{F+\boldsymbol{g}_i}^T$ are block diagonal. Since the addition of $\boldsymbol{g}_i$ can be thought of as a 1–1 mapping of the levels in an attribute of F to the levels of the same attribute in $F+\boldsymbol{g}_i$, this property also holds across the rows corresponding to different attributes in B_F and $B_{F+\boldsymbol{g}_i}$. Thus $B_F B_{F+\boldsymbol{g}_i}^T$ and $B_{F+\boldsymbol{g}_i} B_F^T$ will also be block diagonal.

We then use the expressions for the contrast matrices derived earlier to simplify each of these products. Firstly, $B_F B_F^T$ becomes

$$\begin{aligned} B_F B_F^T &= \text{BlkDiag}\Big[\big(B_3 \otimes \frac{1}{\sqrt{3}}\boldsymbol{j}_3^T\big)\big(B_3^T \otimes \frac{1}{\sqrt{3}}\boldsymbol{j}_3\big), \big(\frac{1}{\sqrt{3}}\boldsymbol{j}_3^T \otimes B_3\big)\big(\frac{1}{\sqrt{3}}\boldsymbol{j}_3 \otimes B_3^T\big), \\ &\qquad \frac{1}{3}\big(B_3Q_0Q_0B_3^T + B_3Q_1Q_{-1}B_3^T + B_3Q_2Q_{-2}B_3^T\big), \\ &\qquad \frac{1}{3}\big(B_3Q_0Q_0B_3^T + B_3Q_2Q_{-2}B_3^T + B_3Q_1Q_{-1}B_3^T\big)\Big] \\ &= \text{BlkDiag}\left[B_3B_3^T, B_3B_3^T, B_3B_3^T, B_3B_3^T\right]. \end{aligned}$$

We can find a similar expression for $B_F B_{F+\boldsymbol{g}_i}^T$:

$$\begin{aligned} B_F B_{F+\boldsymbol{g}_i}^T &= \text{BlkDiag}\Big[\big(B_3 \otimes \frac{1}{\sqrt{3}}\boldsymbol{j}_3^T\big)\big(Q_{-g_{2,1}}B_3^T \otimes \frac{1}{\sqrt{3}}\boldsymbol{j}_3\big), \big(\frac{1}{\sqrt{3}}\boldsymbol{j}_3^T \otimes B_3\big)\big(\frac{1}{\sqrt{3}}\boldsymbol{j}_3 \otimes Q_{-g_{2,2}}B_3^T\big), \\ &\qquad \frac{1}{3}\big(B_3Q_0Q_0Q_{-g_{2,3}}B_3^T + B_3Q_1Q_{-1}Q_{-g_{2,3}}B_3^T + B_3Q_2Q_{-2}Q_{-g_{2,3}}B_3^T\big), \\ &\qquad \frac{1}{3}\big(B_3Q_0Q_0Q_{-g_{2,4}}B_3^T + B_3Q_2Q_{-2}Q_{-g_{2,4}}B_3^T + B_3Q_1Q_{-1}Q_{-g_{2,4}}B_3^T\big)\Big] \\ &= \text{BlkDiag}\left[B_3Q_{-g_{2,1}}B_3^T, B_3Q_{-g_{2,2}}B_3^T, B_3Q_{-g_{2,3}}B_3^T, B_3Q_{-g_{2,4}}B_3^T\right]. \end{aligned}$$

Similarly, the remaining two products are

$$B_{F+\boldsymbol{g}_i}B_F^T = \text{BlkDiag}\left[B_3Q_{g_{2,1}}B_3^T, B_3Q_{g_{2,2}}B_3^T, B_3Q_{g_{2,3}}B_3^T, B_3Q_{g_{2,4}}B_3^T\right],$$

and

$$\begin{aligned} & B_{F+\boldsymbol{g}_i}B_{F+\boldsymbol{g}_i}^T \\ &= \text{BlkDiag}\left[B_3Q_{g_{2,1}}Q_{-g_{2,1}}B_3^T, B_3Q_{g_{2,2}}Q_{-g_{2,2}}B_3^T, B_3Q_{g_{2,3}}Q_{-g_{2,3}}B_3^T, B_3Q_{g_{2,4}}Q_{-g_{2,4}}B_3^T\right] \\ &= \text{BlkDiag}\left[B_3B_3^T, B_3B_3^T, B_3B_3^T, B_3B_3^T\right]. \end{aligned}$$

Then the information matrix for the estimation of main effects only for the first two options across all of the choice sets is

$$\begin{aligned} C(\boldsymbol{\pi}_0)_{1,2} &= \frac{1}{3^2 \times 3^4}\text{BlkDiag}\left[B_3(2I_3 - Q_{g_{2,1}} - Q_{-g_{2,1}})B_3^T, B_3(2I_3 - Q_{g_{2,2}} - Q_{-g_{2,2}})B_3^T,\right. \\ &\qquad \left. B_3(2I_3 - Q_{g_{2,3}} - Q_{-g_{2,3}})B_3^T, B_3(2I_3 - Q_{g_{2,4}} - Q_{-g_{2,4}})B_3^T\right] \\ &= \frac{1}{3^2 \times 3^4}\text{BlkDiag}\left[B_3(2I_3 - Q_1 - Q_2)B_3^T, B_3(2I_3 - Q_1 - Q_2)B_3^T,\right. \\ &\qquad \left. B_3(2I_3 - Q_1 - Q_2)B_3^T, B_3(2I_3 - Q_1 - Q_2)B_3^T\right] \\ &= \frac{1}{3^6}\text{BlkDiag}\left[B_3(3I_3 - J_3)B_3^T, B_3(3I_3 - J_3)B_3^T, B_3(3I_3 - J_3)B_3^T, B_3(3I_3 - J_3)B_3^T\right] \\ &= \frac{1}{3^6}\text{BlkDiag}\left[3B_3B_3^T, 3B_3B_3^T, 3B_3B_3^T, 3B_3B_3^T\right] \\ &= \frac{1}{243}I_8, \end{aligned}$$

since $\sum_{i=0}^{2} Q_i = J_3$.

Of course this approach gives the same value for $C(\boldsymbol{\pi}_0)_{1,2}$ as the numerical calculation. Similarly, the application of this result will give $C(\boldsymbol{\pi}_0)_{1,3} = \frac{1}{243}I_8$ and $C(\boldsymbol{\pi}_0)_{2,3} = \frac{1}{243}I_8$, once again resulting in $C(\boldsymbol{\pi}_0)_M = \frac{1}{81}I_8$. □

This strategy generalises. In general, the information matrix for the pair of items in the first and i^{th} positions is

$$
\begin{aligned}
&m^2NC(\boldsymbol{\pi}_0)_{1i}\\
&= \frac{1}{\ell^{p/2}}\begin{bmatrix} B_{F,1} & B_{F+\boldsymbol{g}_i,1} & B_{\bar{A},1} \\ B_{F,2} & B_{F+\boldsymbol{g}_i,2} & B_{\bar{A},2}\end{bmatrix} \times \frac{1}{m^2\ell^{k-p}}\begin{bmatrix} I_{\ell^{k-p}} & -I_{\ell^{k-p}} & \mathbf{0} \\ -I_{\ell^{k-p}} & I_{\ell^{k-p}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0}\end{bmatrix} \times \frac{1}{\ell^{p/2}}\begin{bmatrix} B^T_{F,1} & B^T_{F,2} \\ B^T_{F+\boldsymbol{g}_i,1} & B^T_{F+\boldsymbol{g}_i,2} \\ B^T_{\bar{A},1} & B^T_{\bar{A},2}\end{bmatrix}\\
&= \frac{1}{m^2\ell^k}\begin{bmatrix} B_{F,1}B^T_{F,1} - B_{F+\boldsymbol{g}_i,1}B^T_{F,1} - B_{F,1}B^T_{F+\boldsymbol{g}_i,1} + B_{F+\boldsymbol{g}_i,1}B^T_{F+\boldsymbol{g}_i,1}, \\ \qquad B_{F,1}B^T_{F,2} - B_{F+\boldsymbol{g}_i,1}B^T_{F,2} - B_{F,1}B^T_{F+\boldsymbol{g}_i,2} + B_{F+\boldsymbol{g}_i,1}B^T_{F+\boldsymbol{g}_i,2} \\ \\ B_{F,2}B^T_{F,1} - B_{F+\boldsymbol{g}_i,2}B^T_{F,1} - B_{F,2}B^T_{F+\boldsymbol{g}_i,1} + B_{F+\boldsymbol{g}_i,2}, B^T_{F+\boldsymbol{g}_i,1}, \\ \qquad B_{F,2}B^T_{F,2} - B_{F+\boldsymbol{g}_i,2}B^T_{F,2} - B_{F,2}B^T_{F+\boldsymbol{g}_i,2} + B_{F+\boldsymbol{g}_i,2}B^T_{F+\boldsymbol{g}_i,2}\end{bmatrix},
\end{aligned}
$$

where $B_{F,1}$ and $B_{F+\boldsymbol{g}_i,1}$ contain the rows of the contrast matrices B_F and $B_{F+\boldsymbol{g}_i}$ respectively corresponding to the first $k-p$ attributes, and $B_{F,2}$ and $B_{F+\boldsymbol{g}_i,2}$ contain the rows of the contrast matrices B_F and $B_{F+\boldsymbol{g}_i}$ respectively corresponding to the remaining p attributes.

By considering each of these matrices individually we simplify this expression for $m^2NC(\boldsymbol{\pi}_0)_{i,j}$ greatly. First, recall from Section 1.B that two contrasts with coefficients λ_i and μ_i are orthogonal if and only if

$$\sum_{i=1}^{\ell}\frac{\lambda_i\mu_i}{n_i} = 0.$$

By construction, each pair of attributes contains all possible pairs of levels equally often. Therefore the product

$$B_{F,(q_1)}B^T_{F,(q_2)},$$

where $B_{F,(q_1)}$ and $B^T_{F,(q_2)}$ are the rows of B_F corresponding to the levels of distinct attributes q_1 and q_2, will contain each combination of λ_i and μ_i equally often and thus

$$B_{F,(q_1)}B^T_{F,(q_2)} = \mathbf{0}.$$

If we add the generator $\boldsymbol{g}_i$ to a starting design, we provide a 1–1 mapping of levels of the attributes in F to the same set of levels in each of the attributes. This preserves the property that within each pair of attributes all of the possible pairs of levels occur equally often. Then

$$B_{F+\boldsymbol{g}_i,(q_1)}B^T_{F+\boldsymbol{g}_i,(q_2)} = \mathbf{0}$$

for $q_1 \neq q_2$. Hence we have

$$
\begin{aligned}
B_{F,1}B^T_{F,2} &= \mathbf{0}, & B_{F+\boldsymbol{g}_i,1}B^T_{F+\boldsymbol{g}_i,2} &= \mathbf{0}, \text{ and}\\
B_{F,2}B^T_{F,1} &= \mathbf{0}, & B_{F+\boldsymbol{g}_i,2}B^T_{F+\boldsymbol{g}_i,1} &= \mathbf{0}.
\end{aligned}
$$

In fact, since the generator forms a 1–1 mapping of the set of attribute levels onto itself, all of the pairs of levels in attribute q_1 of F and attribute q_2 of $F+\boldsymbol{g}_i$ will also occur equally often, for $q_1 \neq q_2$. Then for any generator $\boldsymbol{g}_i$

$$
\begin{aligned}
B_{F,1}B^T_{F+\boldsymbol{g}_i,2} &= \mathbf{0}, & B_{F+\boldsymbol{g}_i,1}B^T_{F,2} &= \mathbf{0}, \text{ and}\\
B_{F,2}B^T_{F+\boldsymbol{g}_i,1} &= \mathbf{0}, & B_{F+\boldsymbol{g}_i,2}B^T_{F,1} &= \mathbf{0}.
\end{aligned}
$$

The same reasoning shows that the matrices

$$\begin{array}{llll} B_{F,1}B_{F,1}^T, & B_{F,1}B_{F+\boldsymbol{g}_i,1}^T, & B_{F+\boldsymbol{g}_i,1}B_{F,1}^T, & B_{F+\boldsymbol{g}_i,1}B_{F+\boldsymbol{g}_i,1}^T, \\ B_{F,2}B_{F,2}^T, & B_{F,2}B_{F+\boldsymbol{g}_i,2}^T, & B_{F+\boldsymbol{g}_i,2}B_{F,2}^T, & B_{F+\boldsymbol{g}_i,2}B_{F+\boldsymbol{g}_i,2}^T \end{array}$$

are block diagonal. For example,

$$\begin{aligned} B_{F,1}B_{F+\boldsymbol{g}_1,1}^T &= \begin{bmatrix} B_{F,(1)} \\ B_{F,(2)} \\ \vdots \\ B_{F,(k-p)} \end{bmatrix} \times \begin{bmatrix} B_{F+\boldsymbol{g}_i,(1)}^T & B_{F+\boldsymbol{g}_i,(2)}^T & \cdots & B_{F+\boldsymbol{g}_i,(k-p)}^T \end{bmatrix} \\ &= \begin{bmatrix} B_{F,(1)}B_{F+\boldsymbol{g}_i,(1)}^T & B_{F,(1)}B_{F+\boldsymbol{g}_i,(2)}^T & \cdots & B_{F,(1)}B_{F+\boldsymbol{g}_i,(k-p)}^T \\ B_{F,(2)}B_{F+\boldsymbol{g}_i,(1)}^T & B_{F,(2)}B_{F+\boldsymbol{g}_i,(2)}^T & \cdots & B_{F,(2)}B_{F+\boldsymbol{g}_i,(k-p)}^T \\ \vdots & \vdots & \ddots & \vdots \\ B_{F,(k-p)}B_{F+\boldsymbol{g}_i,(1)}^T & B_{F,(k-p)}B_{F+\boldsymbol{g}_i,(2)}^T & \cdots & B_{F,(k-p)}B_{F+\boldsymbol{g}_i,(k-p)}^T \end{bmatrix} \\ &= \begin{bmatrix} B_{F,(1)}B_{F+\boldsymbol{g}_i,(1)}^T & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & B_{F,(2)}B_{F+\boldsymbol{g}_i,(2)}^T & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & B_{F,(k-p)}B_{F+\boldsymbol{g}_i,(k-p)}^T \end{bmatrix}. \end{aligned}$$

Using this, we can now find $B_{F,1}B_{F+\boldsymbol{g}_i,1}^T$. Since we know that this matrix is block diagonal, we only need to consider the $(q,q)^{\text{th}}$ block, where $1 \le q \le k-p$. So the row of block matrices of the contrast matrix $B_{F+\boldsymbol{g}_i,1}$ corresponding to the q^{th} attribute will be

$$B_{F+\boldsymbol{g}_i,(q)} = \left[\frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes \ldots \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes (B_\ell Q_{g_{i,q}}) \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes \ldots \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T\right]$$

and the transpose of this row of block matrices will be

$$B_{F+\boldsymbol{g}_i,(q)}^T = \left[\frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell \otimes \ldots \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell \otimes (Q_{-g_{i,q}} B_\ell^T) \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell \otimes \ldots \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell\right].$$

Then the $(q,q)^{\text{th}}$ block of $B_{F,1}B_{F+\boldsymbol{g}_i,1}^T$ will be

$$\begin{aligned} B_{F,(q)}B_{F+\boldsymbol{g}_i,(q)}^T &= \frac{1}{\ell}\boldsymbol{j}_\ell^T\boldsymbol{j}_\ell \otimes \ldots \otimes \frac{1}{\ell}\boldsymbol{j}_\ell^T\boldsymbol{j}_\ell \otimes (B_\ell Q_{-\boldsymbol{g}_{i,q}} B_\ell^T) \otimes \frac{1}{\ell}\boldsymbol{j}_\ell^T\boldsymbol{j}_\ell \otimes \ldots \otimes \frac{1}{\ell}\boldsymbol{j}_\ell^T\boldsymbol{j}_\ell \\ &= 1 \otimes \ldots \otimes 1 \otimes (B_\ell Q_{-g_{i,q}} B_\ell^T) \otimes 1 \otimes \ldots \otimes 1 \\ &= B_\ell Q_{-g_{i,q}} B_\ell^T. \end{aligned}$$

Notice that if $i = 1$ then $Q_{g_{1,q}-g_{i,q}} = I_\ell$, which gives

$$B_{F(q)}B_{F(q)}^T = B_\ell B_\ell^T.$$

Then the $(1,1)$ block of $m^2 NC(\boldsymbol{\pi}_0)_{1i}$ can be simplified to

$$\begin{aligned} \frac{1}{\ell^p}\text{BlkDiag}\big[&B_\ell(2I_{\ell-1} - Q_{g_{i,1}} - Q_{-g_{i,1}})B_\ell^T, B_\ell(2I_{\ell-1} - Q_{g_{i,2}} - Q_{-g_{i,2}})B_\ell^T, \ldots, \\ &B_\ell(2I_{\ell-1} - Q_{g_{i,k-p}} - Q_{-g_{i,k-p}})B_\ell^T\big]. \end{aligned}$$

Now consider the terms involving $B_{F+\boldsymbol{g}_i,2}$. Recall that the q^{th} row of block matrices in $B_{F+\boldsymbol{g}_i,2}$ will be

$$B_{F+\boldsymbol{g}_i,(q)} = \frac{1}{\ell^{(h-1)/2}} \times \begin{bmatrix} (B_\ell Q_{g_i,q} Q_{(0\times a_{q,1}+\ldots+0\times a_{q,h-1})}) \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes \ldots \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T, \\ (B_\ell Q_{g_i,q} Q_{(0\times a_{q,1}+\ldots+0\times a_{q,h-2}+1\times a_{q,h-1})}) \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes \ldots \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T, \ldots, \\ (B_\ell Q_{g_i,q} Q_{(0\times a_{q,1}+\ldots+0\times a_{q,k-2}+\alpha^{\ell-2}\times a_{q,h-1})}) \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes \ldots \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T, \ldots, \\ (B_\ell Q_{g_i,q} Q_{(\alpha^{\ell-2}\times a_{q,1}+\ldots+\alpha^{\ell-2}\times a_{q,h-1})}) \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \otimes \ldots \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell^T \end{bmatrix}$$

where $1 \leq h \leq k-p$ and there will be $(h-1)$ $\boldsymbol{j}_\ell^T$s in each block matrix. Then we may express $B_{F+\boldsymbol{g}_i,2}$ as

$$B_{F+\boldsymbol{g}_i,2} = \begin{bmatrix} B_{F+\boldsymbol{g}_i,(k-p+1)} \\ B_{F+\boldsymbol{g}_i,(k-p+2)} \\ \vdots \\ B_{F+\boldsymbol{g}_i,(k-1)} \\ B_{F+\boldsymbol{g}_i,(k)} \end{bmatrix}$$

and the transpose of $B_{F+g_i,2}$ as

$$B_{F+\boldsymbol{g}_i,2}^T = \begin{bmatrix} B_{F+\boldsymbol{g}_i,(k-p+1)}^T & B_{F+\boldsymbol{g}_i,(k-p+2)}^T & \cdots & B_{F+\boldsymbol{g}_i,(k-1)}^T & B_{F+\boldsymbol{g}_i,(k)}^T \end{bmatrix},$$

where

$$B_{F+g_i,(q)}^T = \frac{1}{\ell^{(h-1)/2}} \times \begin{bmatrix} \left(Q_{-(0\times a_{q,h-1}+\ldots+0\times a_{q,1})} Q_{-g_{i,q}} B_\ell^T\right) \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell \otimes \ldots \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell \\ \left(Q_{-(1\times a_{q,h-1}+0\times a_{q,h-2}+\ldots+0\times a_{q,1})} Q_{-g_{i,q}} B_\ell^T\right) \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell \otimes \ldots \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell \\ \vdots \\ \left(Q_{-(\alpha^{\ell-2}\times a_{q,h-1}+\ldots+\alpha^{\ell-2}\times a_{q,1})} Q_{-g_{i,q}} B_\ell^T\right) \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell \otimes \ldots \otimes \frac{1}{\sqrt{\ell}}\boldsymbol{j}_\ell \end{bmatrix},$$

for $k-p+1 \leq q \leq k$.

We derive the block diagonal entries of $B_{F+\boldsymbol{g}_i,2}B_{F+\boldsymbol{g}_i,2}^T$, by finding the product of the row of block matrices corresponding to the main effects of attribute q and the transpose of same, giving

$$\begin{aligned} B_{F,(q)}B_{F+\boldsymbol{g}_i,(q)}^T &= \frac{1}{\ell^{h-1}} \sum_{b_1=0}^{\ell-1} \cdots \sum_{b_{h-1}=0}^{\ell-1} \left(B_\ell Q_{\sum_{r=1}^{h-1} b_r a_{q,r}} \times Q_{\sum_{r=1}^{h-1} -b_r a_{q,r}} \times Q_{-g_{i,q}} B_\ell^T\right) \\ &= \frac{1}{\ell^{h-1}} \sum_{b_1=0}^{\ell-1} \cdots \sum_{b_{h-1}=0}^{\ell-1} \left(B_\ell Q_{\sum_{r=1}^{h-1}(b_r a_{q,r} - b_r a_{q,r})} \times Q_{-g_{i,q}} B_\ell^T\right) \\ &= \frac{1}{\ell^{h-1}} \sum_{b_1=0}^{\ell-1} \cdots \sum_{b_{h-1}=0}^{\ell-1} B_\ell Q_0 Q_{-g_{i,q}} B_\ell^T \\ &= \frac{\ell^{h-1}}{\ell^{h-1}} B_\ell Q_{-g_{i,q}} B_\ell^T \\ &= B_\ell Q_{-g_{i,q}} B_\ell^T, \end{aligned}$$

and

$$B_{F+\boldsymbol{g}_i,(q)} B_{F,(q)}^T = B_\ell Q_{g_{i,q}} B_\ell^T.$$

The same argument gives

$$\begin{aligned} B_{F,(q)} B_{F,(q)}^T &= B_\ell B_\ell^T, \text{ and} \\ B_{F+\boldsymbol{g}_i,(q)} B_{F+\boldsymbol{g}_i,(q)}^T &= B_\ell Q_0 B_\ell^T \\ &= B_\ell B_\ell^T. \end{aligned}$$

for $k-p+1 \le q \le k$. Putting these results together, we obtain

$$\begin{aligned} m^2 NC(\boldsymbol{\pi}_0)_{1i} = & \frac{1}{\ell^p} \text{BlkDiag}\Big[B_F B_F^T + B_{F+\boldsymbol{g}_{i,1}} B_{F+\boldsymbol{g}_{i,1}}^T - B_F B_{F+\boldsymbol{g}_{i,1}}^T - B_{F+\boldsymbol{g}_{i,1}} B_F^T, \ldots, \\ & B_F B_F^T + B_{F+\boldsymbol{g}_{i,k}} B_{F+\boldsymbol{g}_{i,k}}^T - B_F B_{F+\boldsymbol{g}_{i,k}}^T - B_{F+\boldsymbol{g}_{i,k}} B_F^T \Big] \\ = & \frac{1}{\ell^p} \text{BlkDiag}\left[B_\ell (2I_\ell - Q_{g_{i,1}} - Q_{-g_{i,1}}) B_\ell^T, B_\ell (2I_\ell - Q_{g_{i,2}} - Q_{-g_{i,2}}) B_\ell^T, \ldots, \right. \\ & B_\ell (2I_\ell - Q_{g_{i,k-p}} - Q_{-g_{i,k-p}}) B_\ell^T, B_\ell (2I_\ell - Q_{g_{i,k-p+1}} - Q_{-g_{i,k-p+1}}) B_\ell^T, \\ & \left. B_\ell (2I_\ell - Q_{g_{i,k-p+2}} - Q_{-g_{i,k-p+2}}) B_\ell^T, \ldots, B_\ell (2I_\ell - Q_{g_{i,k}} - Q_{-g_{i,k}}) B_\ell^T \right]. \end{aligned}$$

Then by dividing by $m^2 N$, where $N = \ell^{k-p}$, we obtain the information matrix for the estimation of the main effects of each attribute.

$$C(\boldsymbol{\pi}_0)_{1i} = \frac{1}{m^2 \ell^k} \text{BlkDiag}\left[B_\ell (2I_\ell - Q_{g_{i,1}} - Q_{-g_{i,1}}) B_\ell^T, \ldots, B_\ell (2I_\ell - Q_{g_{i,k}} - Q_{-g_{i,k}}) B_\ell^T \right].$$

We can use a similar argument to show that this is true for any pair of positions, i and j, where $\boldsymbol{g}_j - \boldsymbol{g}_i \notin F$. This expression will be used in Theorem 6.4.1 to obtain optimal designs when all $\boldsymbol{g}_i \notin F$. Now we turn out attention to the case where $\boldsymbol{g}_i \in F$.

6.4 Constructing the Information Matrix - Generator Appears in the Starting Design

When $\boldsymbol{g}_i \in F$, the $\Lambda(\boldsymbol{\pi}_0)$ matrix does not take the same form as when $\boldsymbol{g}_i \notin F$. However, we can consider a different partitioning of the contrast matrix and obtain a useful form for the $\Lambda(\boldsymbol{\pi}_0)$ matrix. We use some of the properties of $GF[\ell]$, where ℓ is a prime power, to do so. We now consider an example of how this will be done.

■ **EXAMPLE 6.4.1.**

Consider a 5^{5-3} experiment with the defining equations

$$\begin{aligned} \boldsymbol{b}_3 &= \boldsymbol{b}_1 + \boldsymbol{b}_2 \\ \boldsymbol{b}_4 &= 2\boldsymbol{b}_1 + \boldsymbol{b}_2 \\ \boldsymbol{b}_5 &= 3\boldsymbol{b}_1 + \boldsymbol{b}_2. \end{aligned}$$

This gives the fractional factorial design shown as the starting design in the first column of Table 6.6. Using the results from Section 6.1, we can express the contrast matrix for the starting design

B_F as

$$B_F = \begin{bmatrix} \multicolumn{5}{c}{B_5 \otimes \frac{1}{\sqrt{5}}\boldsymbol{j}_5^T} \\ \multicolumn{5}{c}{\frac{1}{\sqrt{5}}\boldsymbol{j}_5^T \otimes B_5} \\ B_5 & B_5Q_{1\times 1} & B_5Q_{2\times 1} & B_5Q_{3\times 1} & B_5Q_{4\times 1} \\ B_5 & B_5Q_{1\times 2} & B_5Q_{2\times 2} & B_5Q_{3\times 2} & B_5Q_{4\times 2} \\ B_5 & B_5Q_{1\times 3} & B_5Q_{2\times 3} & B_5Q_{3\times 3} & B_5Q_{4\times 3} \end{bmatrix}$$

$$= \begin{bmatrix} \multicolumn{5}{c}{B_5 \otimes \frac{1}{\sqrt{5}}\boldsymbol{j}_5^T} \\ \multicolumn{5}{c}{\frac{1}{\sqrt{5}}\boldsymbol{j}_5^T \otimes B_5} \\ B_5 & B_5Q_1 & B_5Q_2 & B_5Q_3 & B_5Q_4 \\ B_5 & B_5Q_2 & B_5Q_4 & B_5Q_1 & B_5Q_3 \\ B_5 & B_5Q_3 & B_5Q_1 & B_5Q_4 & B_5Q_2 \end{bmatrix}.$$

To construct the remaining options of each choice set in the choice experiment, we add a set of generators to the starting design. In this example, we will let $m = 3$ and use the generators $\boldsymbol{g}_2 = (11234)$ and $\boldsymbol{g}_3 = (33313)$, noticing that $\boldsymbol{g}_2 \in F$, $\boldsymbol{g}_3 \notin F$, and $\boldsymbol{g}_3 - \boldsymbol{g}_2 = (22134) \notin F$. The design formed by this set of generators is shown in Table 6.6.

In order to calculate the information matrix for the first two options from all of the choice sets, we partition the items in F into five sets of items $\{F_1, F_1+\boldsymbol{g}_2, F_1+2\boldsymbol{g}_2, F_1+3\boldsymbol{g}_2, F_1+4\boldsymbol{g}_2\}$, where F_1 is chosen to make these five sets a partitioning of F. This partitioning is possible because the principal fraction of a fractional factorial design is closed under component–wise addition in $GF[\ell]$. Suppose that we let F_1 consist of those items in F that have the first attribute at the 0 level. Then

$$F_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 2 & 2 & 2 & 2 \\ 0 & 3 & 3 & 3 & 3 \\ 0 & 4 & 4 & 4 & 4 \end{pmatrix},$$

and the remaining sets, shown in Table 6.7, form a partitioning of F. The B_F matrix can then be expressed as

$$B_F = \begin{bmatrix} B_{F_1} & B_{F_1+\boldsymbol{g}_2} & B_{F_1+2\boldsymbol{g}_2} & B_{F_1+3\boldsymbol{g}_2} & B_{F_1+4\boldsymbol{g}_2} \end{bmatrix}.$$

Notice that if $\boldsymbol{g} \in F$ then $\boldsymbol{g}$ and $\boldsymbol{g}+\boldsymbol{g}_2$ appear in the same choice set. Also if $\boldsymbol{g}+4\boldsymbol{g}_2 \in F$ then $\boldsymbol{g}+4\boldsymbol{g}_2$ and $\boldsymbol{g}+4\boldsymbol{g}_2+\boldsymbol{g}_2 = \boldsymbol{g}$ appear together in another choice set. So as written in Table 6.7, each item appears in a choice set with the entry in the column to the left of it and in a choice set with the entry in the column to the right of it, where columns 1 and 5 are viewed as being adjacent. Then $\Lambda(\boldsymbol{\pi}_0)_{1,2}$ is equal to

$$\Lambda(\boldsymbol{\pi}_0)_{1,2} = \frac{1}{m^2N} \begin{bmatrix} 2I_5 & -I_5 & \mathbf{0} & \mathbf{0} & -I_5 & \mathbf{0} \\ -I_5 & 2I_5 & -I_5 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -I_5 & 2I_5 & -I_5 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -I_5 & 2I_5 & -I_5 & \mathbf{0} \\ -I_5 & \mathbf{0} & \mathbf{0} & -I_5 & 2I_5 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

Option 1					Option 2					Option 3				
0	0	0	0	0	1	1	2	3	4	3	3	3	1	3
0	1	1	1	1	1	2	3	4	0	3	4	4	2	4
0	2	2	2	2	1	3	4	0	1	3	0	0	3	0
0	3	3	3	3	1	4	0	1	2	3	1	1	4	1
0	4	4	4	4	1	0	1	2	3	3	2	2	0	2
1	0	1	2	3	2	1	3	0	2	4	3	4	3	1
1	1	2	3	4	2	2	4	1	3	4	4	0	4	2
1	2	3	4	0	2	3	0	2	4	4	0	1	0	3
1	3	4	0	1	2	4	1	3	0	4	1	2	1	4
1	4	0	1	2	2	0	2	4	1	4	2	3	2	0
2	0	2	4	1	3	1	4	2	0	0	3	0	0	4
2	1	3	0	2	3	2	0	3	1	0	4	1	1	0
2	2	4	1	3	3	3	1	4	2	0	0	2	2	1
2	3	0	2	4	3	4	2	0	3	0	1	3	3	2
2	4	1	3	0	3	0	3	1	4	0	2	4	4	3
3	0	3	1	4	4	1	0	4	3	1	3	1	2	2
3	1	4	2	0	4	2	1	0	4	1	4	2	3	3
3	2	0	3	1	4	3	2	1	0	1	0	3	4	4
3	3	1	4	2	4	4	3	2	1	1	1	4	0	0
3	4	2	0	3	4	0	4	3	2	1	2	0	1	1
4	0	4	3	2	0	1	1	1	1	2	3	2	4	0
4	1	0	4	3	0	2	2	2	2	2	4	3	0	1
4	2	1	0	4	0	3	3	3	3	2	0	4	1	2
4	3	2	1	0	0	4	4	4	4	2	1	0	2	3
4	4	3	2	1	0	0	0	0	0	2	2	1	3	4

Table 6.6: The choice sets used in Example 6.4.1

Then the information matrix for the first two options in the choice sets is

$$C(\boldsymbol{\pi}_0)_{1,2} = \frac{1}{5^{3/2}} \begin{bmatrix} B_{F_1} & B_{F_1+\boldsymbol{g}_2} & B_{F_1+2\boldsymbol{g}_2} & B_{F_1+3\boldsymbol{g}_2} & B_{F_1+4\boldsymbol{g}_2} & B_{\bar{A}} \end{bmatrix} \times \frac{1}{225} \begin{bmatrix} 2I_5 & -I_5 & \mathbf{0} & \mathbf{0} & -I_5 & \mathbf{0} \\ -I_5 & 2I_5 & -I_5 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -I_5 & 2I_5 & -I_5 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & -I_5 & 2I_5 & -I_5 & \mathbf{0} \\ -I_5 & \mathbf{0} & \mathbf{0} & -I_5 & 2I_5 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix} \times \frac{1}{5^{3/2}} \begin{bmatrix} B_{F_1}^T \\ B_{F_1+\boldsymbol{g}_2}^T \\ B_{F_1+2\boldsymbol{g}_2}^T \\ B_{F_1+3\boldsymbol{g}_2}^T \\ B_{F_1+4\boldsymbol{g}_2}^T \\ B_{\bar{A}}^T \end{bmatrix}$$

$\mathbf{F_1}$	$\mathbf{F_1+g_2}$	$\mathbf{F_1+2g_2}$	$\mathbf{F_1+3g_2}$	$\mathbf{F_1+4g_2}$
0 0 0 0 0	1 1 2 3 4	2 2 4 1 3	3 3 1 4 2	4 4 3 2 1
0 1 1 1 1	1 2 3 4 0	2 3 0 2 4	3 4 2 0 3	4 0 4 3 2
0 2 2 2 2	1 3 4 0 1	2 4 1 3 0	3 0 3 1 4	4 1 0 4 3
0 3 3 3 3	1 4 0 1 2	2 0 2 4 1	3 1 4 2 0	4 2 1 0 4
0 4 4 4 4	1 0 1 2 3	2 1 3 0 2	3 2 0 3 1	4 3 2 1 0

Table 6.7: The partitions of B_F in Example 6.4.1

$$
\begin{aligned}
&C(\boldsymbol{\pi}_0)_{1,2} \\
&= \frac{1}{225\times 5^3} \times \left[\; B_{F_1}\; B_{F_1+\boldsymbol{g}_2}\; B_{F_1+2\boldsymbol{g}_2}\; B_{F_1+3\boldsymbol{g}_2}\;\; B_{F_1+4\boldsymbol{g}_2} \;\right] \\
&\qquad \times \left(2 \begin{bmatrix} B^T_{F_1} \\ B^T_{F_1+\boldsymbol{g}_2} \\ B^T_{F_1+2\boldsymbol{g}_2} \\ B^T_{F_1+3\boldsymbol{g}_2} \\ B^T_{F_1+4\boldsymbol{g}_2} \end{bmatrix} - \begin{bmatrix} B^T_{F_1+\boldsymbol{g}_2} \\ B^T_{F_1+2\boldsymbol{g}_2} \\ B^T_{F_1+3\boldsymbol{g}_2} \\ B^T_{F_1+4\boldsymbol{g}_2} \\ B^T_{F_1} \end{bmatrix} - \begin{bmatrix} B^T_{F_1+4\boldsymbol{g}_2} \\ B^T_{F_1} \\ B^T_{F_1+\boldsymbol{g}_2} \\ B^T_{F_1+2\boldsymbol{g}_2} \\ B^T_{F_1+3\boldsymbol{g}_2} \end{bmatrix} \right) \\
&= \frac{1}{225\times 5^3}(2B_F B_F^T - B_F B^T_{F+\boldsymbol{g}_2} - B_F B^T_{F-\boldsymbol{g}_2}) \\
&= \frac{1}{225\times 5^3}\text{BlkDiag}\left[B_5(2I_5 - Q_{g_{2,1}} - Q_{-g_{2,1}})B_5^T, B_5(2I_5 - Q_{g_{2,2}} - Q_{-g_{2,2}})B_5^T,\right. \\
&\qquad B_5(2I_5 - Q_{g_{2,3}} - Q_{-g_{2,3}})B_5^T, B_5(2I_5 - Q_{g_{2,4}} - Q_{-g_{2,4}})B_5^T, \\
&\qquad \left. B_5(2I_5 - Q_{g_{2,5}} - Q_{-g_{2,5}})B_5^T\right] \\
&= \frac{1}{225\times 5^3}\text{BlkDiag}\left[B_5(2I_5 - Q_3 - Q_2)B_5^T, B_5(2I_5 - Q_3 - Q_2)B_5^T, B_5(2I_5 - Q_3 - Q_2)B_5^T,\right. \\
&\qquad \left. B_5(2I_5 - Q_1 - Q_4)B_5^T, B_5(2I_5 - Q_3 - Q_2)B_5^T\right].
\end{aligned}
$$

We can find $C(\boldsymbol{\pi}_0)_{1,3}$ and $C(\boldsymbol{\pi}_0)_{2,3}$ using the method that was shown in Example 6.3.2, which gives

$$
\begin{aligned}
C(\boldsymbol{\pi}_0)_{1,3} &= \frac{1}{225\times 5^3}\text{BlkDiag}[B_5(2I_5 - Q_3 - Q_2)B_5^T, B_5(2I_5 - Q_3 - Q_2)B_5^T, \\
&\qquad B_5(2I_5 - Q_3 - Q_2)B_5^T, B_5(2I_5 - Q_1 - Q_4)B_5^T, \\
&\qquad B_5(2I_5 - Q_3 - Q_2)B_5^T],
\end{aligned}
$$

and

$$
\begin{aligned}
C(\boldsymbol{\pi}_0)_{2,3} &= \frac{1}{225\times 5^3}\text{BlkDiag}[B_5(2I_5 - Q_2 - Q_3)B_5^T, B_5(2I_5 - Q_2 - Q_3)B_5^T, \\
&\qquad B_5(2I_5 - Q_1 - Q_4)B_5^T, B_5(2I_5 - Q_3 - Q_2)B_5^T, \\
&\qquad B_5(2I_5 - Q_4 - Q_1)B_5^T].
\end{aligned}
$$

By adding these matrices and substituting the Q_i, B_5 and I_5 matrices, we obtain the information matrix shown in Table 6.8. □

We now use some of the properties of Galois fields to generalise this example. Since the principal fraction of a regular fractional factorial design is closed under component–wise addition

$$
\frac{1}{196875}\left(\begin{array}{cccccccccccccccccccc}
56 & 0 & -7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 60 & 0 & -\sqrt{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-7 & 0 & 49 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -\sqrt{5} & 0 & 45 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 56 & 0 & -7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 60 & 0 & -\sqrt{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -7 & 0 & 49 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -\sqrt{5} & 0 & 45 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 56 & 0 & -7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 60 & 0 & -\sqrt{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -7 & 0 & 49 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\sqrt{5} & 0 & 45 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 56 & 0 & -7 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 60 & 0 & -\sqrt{5} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -7 & 0 & 49 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\sqrt{5} & 0 & 45 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 49 & 0 & 7 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 45 & 0 & \sqrt{5} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 7 & 0 & 56 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{5} & 0 & 60
\end{array}\right)
$$

Table 6.8: The $C(\boldsymbol{\pi}_0)_M$ matrix for the design in Example 6.4.1

in $GF[\ell]$, by adding a generator that is an item within the fractional factorial design to each row of the fractional factorial design, we will yield a rearrangement of the rows of the fractional factorial design.

Since the entries in the generator are elements of $GF[\ell]$, for a non–zero element of $\boldsymbol{g}_i$, say $g_{i,q}$, the set

$$\{0 \times g_{i,q}, 1 \times g_{i,q}, \ldots, \alpha^{\ell-2} \times g_{i,q}\}$$

contains each element of $GF[\ell]$ exactly once. Then for any generator $\boldsymbol{g}_i \in F$, or difference in generators $\boldsymbol{g}_j - \boldsymbol{g}_i$, where each of the entries in the generator are non–zero, the set of generators

$$G = \{0 \times \boldsymbol{g}_i, 1 \times \boldsymbol{g}_i, \ldots, \alpha^{\ell-2} \times \boldsymbol{g}_i\}$$

will contain each element of $GF[\ell]$ exactly once in each position, and therefore the set of generators will form a subgroup of F.

With a suitable selection of F_1, we partition the items in F to form ℓ distinct sets F_1, $F_1 + \boldsymbol{g}_i$, $F_1 + 2 \times \boldsymbol{g}_i$, …, $F_1 + \alpha^{\ell-2} \times \boldsymbol{g}_i$. We partition the columns in the contrast matrix B_F in a similar way; that is,

$$B_F = \begin{bmatrix} B_{F_1} & B_{F_1+\boldsymbol{g}_i} & B_{F_1+2\times\boldsymbol{g}_i} & \ldots & B_{F_1+\alpha^{\ell-2}\times\boldsymbol{g}_i} \end{bmatrix}.$$

Then if $\boldsymbol{g}_i \in F$ is used as a generator, the $\Lambda(\boldsymbol{\pi}_0)_{1i}$ matrix can be written as

$$m^2 N\Lambda(\boldsymbol{\pi}_0)_{1i}$$
$$= \begin{bmatrix} 2I_{\ell^{k-p-1}} & -I_{\ell^{k-p-1}} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} & \mathbf{0} & -I_{\ell^{k-p-1}} & \mathbf{0} \\ -I_{\ell^{k-p-1}} & 2I_{\ell^{k-p-1}} & -I_{\ell^{k-p-1}} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & -I_{\ell^{k-p-1}} & 2I_{\ell^{k-p-1}} & -I_{\ell^{k-p-1}} & \mathbf{0} \\ -I_{\ell^{k-p-1}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} & -I_{\ell^{k-p-1}} & 2I_{\ell^{k-p-1}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

since each item in F will appear as an option in the same choice set as two other items, one belonging to the partition to the left of it, and one belonging the partition to the right of it when B_F is written in this way. Then by substituting the expression for $\Lambda(\boldsymbol{\pi}_0)_{1i}$ into the identity $C(\boldsymbol{\pi}_0)_{1i} = B\Lambda(\boldsymbol{\pi}_0)_{1i}B^T$, and taking note that items that are not in F will not appear as options in positions 1 and i of the choice sets, we obtain

$$m^2 NB\Lambda(\boldsymbol{\pi}_0)_{1i}B^T$$
$$= \frac{1}{\ell^{p/2}} \left[B_{F_1}, B_{F_1+\boldsymbol{g}_i}, B_{F_1+\alpha\times\boldsymbol{g}_i}, \ldots, B_{F_1+\alpha^{\ell-2}\times\boldsymbol{g}_i}, B_{\bar{A}}\right]$$
$$\times \begin{bmatrix} 2I_{\ell^{k-p-1}} & -I_{\ell^{k-p-1}} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} & \mathbf{0} & -I_{\ell^{k-p-1}} & \mathbf{0} \\ -I_{\ell^{k-p-1}} & 2I_{\ell^{k-p-1}} & -I_{\ell^{k-p-1}} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & -I_{\ell^{k-p-1}} & 2I_{\ell^{k-p-1}} & -I_{\ell^{k-p-1}} & \mathbf{0} \\ -I_{\ell^{k-p-1}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} & -I_{\ell^{k-p-1}} & 2I_{\ell^{k-p-1}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}$$
$$\times \frac{1}{\ell^{p/2}} \left[B_{F_1}, B_{F_1+\boldsymbol{g}_i}, B_{F_1+\alpha\times\boldsymbol{g}_i}, \ldots, B_{F_1+\alpha^{\ell-2}\times\boldsymbol{g}_i}, B_{\bar{A}}\right]^T$$

$$\begin{aligned}
& m^2 NB\Lambda(\boldsymbol{\pi}_0)_{1i}B^T \\
= \quad & \frac{1}{\ell p}\sum_{j=0}^{\alpha^{\ell-2}}\left[B_{F_1+j\times \boldsymbol{g}_i}\left(-B^T_{F_1+(j-1)\times \boldsymbol{g}_i}+2B^T_{F_1+j\times \boldsymbol{g}_i}-B^T_{F_1+(j+1)\times \boldsymbol{g}_i}\right)\right] \\
= \quad & \frac{1}{\ell p}\sum_{j=0}^{\alpha^{\ell-2}}\left[2B_{F_1+j\times \boldsymbol{g}_i}B^T_{F_1+j\times \boldsymbol{g}_i}-B_{F_1+j\times \boldsymbol{g}_i}B^T_{F_1+(j-1)\times \boldsymbol{g}_i}-B_{F_1+j\times \boldsymbol{g}_i}B^T_{F_1+(j+1)\times \boldsymbol{g}_i}\right].
\end{aligned}$$

Since the set $\{F_1 + x \times \boldsymbol{g}_i | 0 \le x \le \alpha^{\ell-2}\}$ is a partition of F, we can simplify the components of $m^2NC(\boldsymbol{\pi}_0)_{1i}$ above to obtain $B_F B_F^T$, $B_{F+\boldsymbol{g}_i}B_F^T$, and $B_F B^T_{F+\boldsymbol{g}_i}$ respectively. Thus, we can use some of the results from Section 6.3 to obtain

$$\begin{aligned}
C(\boldsymbol{\pi}_0)_{1i} &= m^2 NB\Lambda(\boldsymbol{\pi}_0)_{1i}B^T \\
&= \frac{1}{\ell p}\big(-B_F B^T_{F+\boldsymbol{g}_i}+2B_F B_F^T - B_{F+\boldsymbol{g}_i}B_F^T\big) \\
&= \frac{1}{\ell p}\text{BlkDiag}\left[B_\ell(2I_\ell - Q_{g_{i,1}}+Q_{-g_{i,1}})B_\ell^T,\ldots,B_\ell(2I_\ell - Q_{g_{i,k}}+Q_{-g_{i,k}})B_\ell^T\right].
\end{aligned}$$

This is identical to the form of $C(\boldsymbol{\pi}_0)_{1i}$ when $\boldsymbol{g}_i \notin F$, as derived in Section 6.4. A similar argument can be used to show that this result holds for an arbitrary i and j, where $\boldsymbol{g}_j - \boldsymbol{g}_i \in F$. If there is a mix of generators, some $\boldsymbol{g}_i \in F$ and some $\boldsymbol{g}_i \notin F$, then we can treat the information matrices for each pair of attributes identically.

If, however, we choose $m = \ell$ generators such that the generators form a subgroup of F, then we obtain a special case of the above construction that allows us to reduce the number of choice sets while maintaining the efficiency of the design. This property is explored in the next example.

■ **EXAMPLE 6.4.2.**

Let us consider a set of generators that form a subgroup of F. We may modify the fractional factorial design from Example 6.1.1 to ensure that a suitable set of generators can be found. To do this, we must enforce the restriction

$$a_{i,1} + a_{i,2} \neq 0,$$

thus allowing at least one item in F to have non–zero entries for each attribute. Table 6.9 gives one such starting design, with (112) and (221) as items with non–zero entries for each attribute.

F								
0	0	0	0	1	1	0	2	2
1	0	1	1	1	2	1	2	0
2	0	2	2	1	0	2	2	1

Table 6.9: The 3^{3-1} fractional factorial design used in Example 6.4.2

We then choose a subgroup of the fractional factorial design F to form the set of generators. The set $G = (000, 112, 221)$ forms such a subgroup, which gives the design in Table 6.10. We may also partition the items into three groups $\{F_1, F_1 + \boldsymbol{g}_2, F_1 + 2\boldsymbol{g}_2\}$, the three blocks of columns on

Table 6.9. Since $\boldsymbol{g}_3 = 2\boldsymbol{g}_2$, we may write the $\Lambda(\boldsymbol{\pi}_0)$ matrix as

$$\begin{aligned}\Lambda(\boldsymbol{\pi}_0) &= \sum_{i<j}\Lambda(\boldsymbol{\pi}_0)_{i,j} \\ &= 3\times\frac{1}{3^2\times 9}\begin{bmatrix} 2I_3 & -I_3 & -I_3 \\ -I_3 & 2I_3 & -I_3 \\ -I_3 & -I_3 & 2I_3 \end{bmatrix} \\ &= \frac{1}{27}\begin{bmatrix} 2I_3 & -I_3 & -I_3 \\ -I_3 & 2I_3 & -I_3 \\ -I_3 & -I_3 & 2I_3 \end{bmatrix},\end{aligned}$$

where the contrast matrix can be partitioned into the contrast matrix for F_1, and the contrast matrix for F_1 plus multiples of the generator. This gives

$$B = \frac{1}{\sqrt{3}}\begin{bmatrix} B_{F_1} & B_{F_1+\boldsymbol{g}_2} & B_{F_1+2\boldsymbol{g}_2} \end{bmatrix}.$$

It follows that the information matrix for the estimation of main effects is

$$C(\boldsymbol{\pi}_0)_M = \frac{1}{27}I_8.$$

We notice that the $\Lambda(\boldsymbol{\pi}_0)_{i,j}$ matrices are identical for each pair of options. This suggests that there is some scope to decrease the number of choice sets used to obtain this information. We could use F_1 as the starting design and then let each subgroup form a separate option within the choice set, thereby reducing the number of choice sets to the 3 shown in Table 6.11. Then the $\Lambda(\boldsymbol{\pi}_0)_{i,j}$ matrices for each of the pairs of options will become

$$\Lambda(\boldsymbol{\pi}_0)_{1,2} = \frac{1}{3^2\times 3}\begin{bmatrix} I_3 & -I_3 & \mathbf{0} \\ -I_3 & I_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \qquad \Lambda(\boldsymbol{\pi}_0)_{1,3} = \frac{1}{3^2\times 3}\begin{bmatrix} I_3 & \mathbf{0} & -I_3 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -I_3 & \mathbf{0} & I_3 \end{bmatrix},$$

$$\text{and } \Lambda(\boldsymbol{\pi}_0)_{2,3} = \frac{1}{3^2\times 3}\begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I_3 & -I_3 \\ \mathbf{0} & -I_3 & I_3 \end{bmatrix}.$$

Since $\Lambda(\boldsymbol{\pi}_0)$ is the sum of these $\Lambda(\boldsymbol{\pi}_0)_{i,j}$ matrices, we have

$$\begin{aligned}\Lambda(\boldsymbol{\pi}_0) &= \sum_{i<j}\Lambda(\boldsymbol{\pi}_0)_{i,j} \\ &= \frac{1}{3^2\times 3}\begin{bmatrix} 2I_3 & -I_3 & -I_3 \\ -I_3 & 2I_3 & -I_3 \\ -I_3 & -I_3 & 2I_3 \end{bmatrix}.\end{aligned}$$

It follows that the information matrix for the estimation of main effects will be

$$C(\boldsymbol{\pi}_0)_M = \frac{1}{27}I_8,$$

as before. This demonstrates that if the set of generators forms a subgroup of F, there is scope to reduce the number of choice sets without changing the D–efficiency of the design. □

We can generalise the example to other cases where the set of generators form a subgroup of F. That is, the choice sets formed where multiples of $\boldsymbol{g}_i$ are used to generate the $m = \ell$ options

Option 1			Option 2			Option 3		
0	0	0	1	1	2	2	2	1
1	0	1	2	1	0	0	2	2
2	0	2	0	1	1	1	2	0
0	1	1	1	2	0	2	0	2
1	1	2	2	2	1	0	0	0
2	1	0	0	2	2	1	0	1
0	2	2	1	0	1	2	1	0
1	2	0	2	0	2	0	1	1
2	2	1	0	0	0	1	1	2

Table 6.10: The 3^{3-1} choice experiment used in Example 6.4.2 - 9 choice sets

Option 1			Option 2			Option 3		
0	0	0	1	1	2	2	2	1
0	1	1	1	2	0	2	0	2
0	2	2	1	0	1	2	1	0

Table 6.11: The 3^{3-1} choice experiment used in Example 6.4.2 - 3 choice sets

of the choice set. If we begin by using F_1 as the starting design then within each pair of options, each partition will appear with every other partition in each pair of options. Then the $\Lambda(\boldsymbol{\pi}_0)_{i,j}$ matrix is

$$\Lambda(\boldsymbol{\pi}_0)_{i,j} = \frac{1}{m^2N}\begin{bmatrix} (\ell-1)I_{\ell^{k-p-1}} & -I_{\ell^{k-p-1}} & -I_{\ell^{k-p-1}} & \dots & -I_{\ell^{k-p-1}} & \mathbf{0} \\ -I_{\ell^{k-p-1}} & (\ell-1)I_{\ell^{k-p-1}} & -I_{\ell^{k-p-1}} & \dots & -I_{\ell^{k-p-1}} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -I_{\ell^{k-p-1}} & -I_{\ell^{k-p-1}} & -I_{\ell^{k-p-1}} & \dots & (\ell-1)I_{\ell^{k-p-1}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \end{bmatrix}.$$

We now proceed to derive the information matrix $C(\boldsymbol{\pi}_0)$ for the estimation of main effects when using this design. Recall that

$$C(\boldsymbol{\pi}_0)_M = B\Lambda(\boldsymbol{\pi}_0)B^T.$$

The product of the first two matrices is

$$\begin{aligned} B\Lambda(\boldsymbol{\pi}_0) &= \frac{1}{m^2N}\begin{bmatrix} B_{F_1} & B_{F_1+\boldsymbol{g}_i} & B_{F_1+\alpha\times\boldsymbol{g}_i} & \dots & B_{F_1+\alpha^{\ell-2}\times\boldsymbol{g}_i} & B_{\bar{A}} \end{bmatrix} \\ &\times\ell\times\begin{bmatrix} (\ell-1)I_{\ell^{k-p-1}} & -I_{\ell^{k-p-1}} & -I_{\ell^{k-p-1}} & \dots & -I_{\ell^{k-p-1}} & \mathbf{0} \\ -I_{\ell^{k-p-1}} & (\ell-1)I_{\ell^{k-p-1}} & -I_{\ell^{k-p-1}} & \dots & -I_{\ell^{k-p-1}} & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -I_{\ell^{k-p-1}} & -I_{\ell^{k-p-1}} & -I_{\ell^{k-p-1}} & \dots & (\ell-1)I_{\ell^{k-p-1}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \end{bmatrix} \end{aligned}$$

$$B\Lambda(\boldsymbol{\pi}_0) = \frac{\ell}{m^2N}\Bigg[\ell B_{F_1} - \sum_{j=0}^{\alpha^{\ell-2}} B_{F_1+j\times\boldsymbol{g}_i}, \ell B_{F_1+\boldsymbol{g}_i} - \sum_{j=0}^{\alpha^{\ell-2}} B_{F_1+j\times\boldsymbol{g}_i},$$
$$\ell B_{F_1+\alpha\times\boldsymbol{g}_i} - \sum_{j=0}^{\alpha^{\ell-2}} B_{F_1+j\times\boldsymbol{g}_i}, \ldots, \ell B_{F_1+\alpha^{\ell-2}\times\boldsymbol{g}_i} - \sum_{j=0}^{\alpha^{\ell-2}} B_{F_1+j\times\boldsymbol{g}_i}, \mathbf{0}\Bigg].$$

However, since each element of $\boldsymbol{g}_i$ is non-zero and the addition and multiplication is performed within $GF[\ell]$, then $0\times\boldsymbol{g}_i, 1\times\boldsymbol{g}_i, \ldots, \alpha^{\ell-2}\times\boldsymbol{g}_i$, will contain each element of $GF[\ell]$ exactly once in each position. Then each entry in $\sum_{j=0}^{\alpha^{\ell-2}} B_{F_1+j\times\boldsymbol{g}_i}$ will be the sum of the corresponding row of B_ℓ, giving

$$\sum_{j=0}^{\alpha^{\ell-2}} B_{F_1+j\times\boldsymbol{g}_i} = \mathbf{0},$$

and therefore $B\Lambda(\boldsymbol{\pi}_0)$ can be simplified to

$$B\Lambda(\boldsymbol{\pi}_0) = \frac{\ell}{m^2N}\begin{bmatrix}\ell B_{F_1} & \ell B_{F_1+\boldsymbol{g}_i} & \ell B_{F_1+\alpha\times\boldsymbol{g}_i} & \ldots & \ell B_{F_1+\alpha^{\ell-2}\times\boldsymbol{g}_i} & \mathbf{0}\end{bmatrix}.$$

Then by post–multiplying $B\Lambda(\boldsymbol{\pi}_0)$ by B^T, we obtain

$$\begin{aligned}
B\Lambda(\boldsymbol{\pi}_0)B^T &= \frac{\ell}{m^2N}\begin{bmatrix}\ell B_{F_1} & \ell B_{F_1+\boldsymbol{g}_i} & \ell B_{F_1+\alpha\times\boldsymbol{g}_i} & \ldots & \ell B_{F_1+\alpha^{\ell-2}\times\boldsymbol{g}_i} & \mathbf{0}\end{bmatrix}\\
&\qquad\times\begin{bmatrix}B_{F_1} & B_{F_1+\boldsymbol{g}_i} & B_{F_1+\alpha\times\boldsymbol{g}_i} & \ldots & B_{F_1+\alpha^{\ell-2}\times\boldsymbol{g}_i} & B_{\bar{A}}^T\end{bmatrix}^T\\
&= \frac{1}{m^2N}\ell^2 B_F \times B_F^T\\
&= \frac{1}{m^2N}\ell^2 I_{k(\ell-1)}.
\end{aligned}$$

Then the information matrix for the estimation of main effects will be

$$\begin{aligned}
C(\boldsymbol{\pi}_0)_M &= \frac{1}{\ell^p}B\Lambda(\boldsymbol{\pi}_0)B^T\\
&= \frac{1}{m^2N}\frac{1}{\ell^p}\ell^2 I_{k(\ell-1)}\\
&= \frac{1}{\ell^k}I_{k(\ell-1)}.
\end{aligned}$$

The determinant of this information matrix is

$$\det(C(\boldsymbol{\pi}_0)_M) = \left(\frac{1}{\ell^k}\right)^{k(\ell-1)}.$$

While this would yield an efficient design based on the determinant of the information matrix $C(\boldsymbol{\pi}_0)_M$, we can improve on this design as we did in Example 6.4.2. If, instead of using F as the starting design we use F_1, then we find that the choice sets will be

$$\begin{bmatrix}F_1 & F_1+\boldsymbol{g}_i & F_1+\alpha\times\boldsymbol{g}_i & \ldots & F_1+\alpha^{\ell-2}\times\boldsymbol{g}_i\end{bmatrix}.$$

We notice that each subgroup of F still appears with every other subgroup of F, but in this design each pair appears only once, and not in each pair of options as was the case in the previous design. Therefore

$$\Lambda(\boldsymbol{\pi}_0) = \frac{1}{m^2N}\begin{bmatrix}
(\ell-1)I_{\ell^{k-p-1}} & -I_{\ell^{k-p-1}} & -I_{\ell^{k-p-1}} & \ldots & -I_{\ell^{k-p-1}} & \mathbf{0}\\
-I_{\ell^{k-p-1}} & (\ell-1)I_{\ell^{k-p-1}} & -I_{\ell^{k-p-1}} & \ldots & -I_{\ell^{k-p-1}} & \mathbf{0}\\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\
-I_{\ell^{k-p-1}} & -I_{\ell^{k-p-1}} & -I_{\ell^{k-p-1}} & \ldots & (\ell-1)I_{\ell^{k-p-1}} & \mathbf{0}\\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} & \mathbf{0}
\end{bmatrix},$$

and it follows that

$$\begin{aligned} B\Lambda(\boldsymbol{\pi}_0)B^T &= \frac{1}{m^2N}\left[\ell B_{F_1} \quad \ell B_{F_1+\boldsymbol{g}_i} \quad \ell B_{F_1+\alpha\times\boldsymbol{g}_i} \quad \cdots \quad \ell B_{F_1+\alpha^{\ell-2}\times\boldsymbol{g}_i} \quad \mathbf{0}\right] \\ &\quad \times\left[\ell B_{F_1} \quad \ell B_{F_1+\boldsymbol{g}_i} \quad \ell B_{F_1+\alpha\times\boldsymbol{g}_i} \quad \cdots \quad \ell B_{F_1+\alpha^{\ell-2}\times\boldsymbol{g}_i} \quad B_A^T\right]^T \\ &= \frac{1}{m^2N}\ell B_F \times B_F^T \\ &= \frac{1}{m^2N}\ell I_{k(\ell-1)}. \end{aligned}$$

This design will contain only $N = \ell^{k-p-1}$ choice sets since the subset F_1 was used as the starting design. Then the information matrix will be

$$\begin{aligned} C(\boldsymbol{\pi}_0)_M &= \frac{1}{m^2N}\frac{1}{\ell^p}\ell I_{k(\ell-1)} \\ &= \frac{1}{\ell^k}I_{k(\ell-1)}, \end{aligned}$$

which gives the same D–efficiency as the previous design, but uses fewer choice sets to do so.

The expressions for $C(\boldsymbol{\pi}_0)_M$ derived in the last two sections can be used to find optimal designs for choice experiments where only main effects are of interest. Theorem 6.4.1 establishes the conditions under which the design will be optimal.

■ **THEOREM 6.4.1.**

Let F be the principal block of a regular fractional factorial design for k attributes with ℓ levels each, where ℓ is a prime power. Suppose that we choose a collection of sets of generators $G_\alpha = \{\boldsymbol{g}_{\alpha,1} = \mathbf{0}, \boldsymbol{g}_{\alpha,2}, \ldots, \boldsymbol{g}_{\alpha,m}\}$, $\alpha = 1, \ldots, \zeta$, such that $\boldsymbol{g}_{\alpha,i} \neq \boldsymbol{g}_{\alpha,j}$ for all $i \neq j$. Let $\boldsymbol{g}_{\alpha,i} = (g_{\alpha,i,1}, g_{\alpha,i,2}, \ldots, g_{\alpha,i,k})$ for $i = 1, 2, \ldots, m$ and also suppose that the multiset of differences for each attribute over all sets in the collection

$$\mathop{\&}_{\alpha=1}^{\zeta} \mathop{\&}_{q_1 \neq q_2} (g_{\alpha,q_1,i} - g_{\alpha,q_2,i})$$

contains each non–zero difference in $GF[\ell]$ equally often. Then the choice sets given by the rows of $F + \boldsymbol{g}_{\alpha,1}, F + \boldsymbol{g}_{\alpha,2}, \ldots, F + \boldsymbol{g}_{\alpha,m}$, for $\alpha = 1, \ldots, \zeta$, are optimal for the estimation of main effects only, provided that there are as few zero differences as possible. □

Proof. Suppose that there are ζ sets of generators. Then there will be $\zeta \times \ell^{k-p}$ choice sets in total. Since information matrices are additive we may consider each set of generators G_α separately, and add the information matrices.

Also let $m = \ell x + y$, where there are y entries that are repeated $x + 1$ times in the choice set and $\ell - y$ entries repeated x times in the choice set. Then if we use the assumption that each difference will appear equally often, there will be

$$\zeta\left(m(m-1) - (x+1)xy - x(x-1)(\ell-1)\right) = 2S_q$$

non–zero differences, $\frac{2}{\ell-1}S_q$ of each type of non–zero difference. The remaining

$$\zeta\left((x+1)xy + x(x-1)(\ell-1)\right) = \zeta m(m-1) - 2S_q$$

differences in the difference vector will be zero differences. Then, by using the assumption that the multiset of differences for each attribute i,

$$\mathop{\&}_{\alpha=1}^{\zeta} \mathop{\&}_{q_1 \neq q_2} (g_{\alpha,q_1,i} - g_{\alpha,q_2,i}),$$

contains each non–zero difference in $GF[\ell]$ equally often, and that $Q_0 = I_\ell$ and $\sum_{i=0}^{\alpha^{\ell-2}} Q_i = J_\ell$, the sum of the permutation matrices can be expressed as

$$\begin{aligned}\sum_{i\neq j} Q_{g_{j,r}-g_{i,r}} &= \frac{2}{\ell-1}S_q(J_\ell - I_\ell) + (\zeta m(m-1) - 2S_q)I_\ell \\ &= \frac{2}{\ell-1}S_q J_\ell + \left(\zeta m(m-1) - \frac{2\ell S_q}{\ell-1}\right) I_\ell.\end{aligned}$$

Substituting this into the expression for $C(\boldsymbol{\pi}_0)_{i,j}$ derived earlier, we obtain

$$\begin{aligned}&\sum_{\alpha=1}^{\zeta}\sum_{i<j}(C(\boldsymbol{\pi}_0)_\alpha)_{i,j} \\ &= \frac{1}{m^2\zeta\ell^k}\text{BlkDiag}\Bigg[B_\ell\Bigg(\frac{\zeta m(m-1)}{2}\times 2I_\ell - \sum_{i<j}Q_{g_{i,1}-g_{j,1}} - \sum_{i<j}Q_{g_{j,1}-g_{i,1}}\Bigg)B_\ell^T, \\ &\qquad B_\ell\Bigg(\frac{\zeta m(m-1)}{2}\times 2I_\ell - \sum_{i<j}Q_{g_{i,2}-g_{j,2}} - \sum_{i<j}Q_{g_{j,2}-g_{i,2}}\Bigg)B_\ell^T, \\ &\qquad \ldots, B_\ell\Bigg(\frac{\zeta m(m-1)}{2}\times 2I_\ell - \sum_{i<j}Q_{g_{i,k}-g_{j,k}} - \sum_{i<j}Q_{g_{j,k}-g_{i,k}}\Bigg)B_\ell^T\Bigg] \\ &= \frac{1}{m^2\zeta\ell^k}\text{BlkDiag}\Bigg[B_\ell\Bigg(\zeta m(m-1)I_\ell - \frac{2}{\ell-1}S_qJ_\ell + \left(\zeta m(m-1) - \frac{2\ell S_q}{\ell-1}\right)I_\ell\Bigg)B_\ell^T, \\ &\qquad B_\ell\Bigg(\zeta m(m-1)I_\ell - \frac{2}{\ell-1}S_qJ_\ell + \left(\zeta m(m-1) - \frac{2\ell S_q}{\ell-1}\right)I_\ell\Bigg)B_\ell^T, \\ &\qquad \ldots, B_\ell\Bigg(\zeta m(m-1)I_\ell - \frac{2}{\ell-1}S_qJ_\ell + \left(\zeta m(m-1) - \frac{2\ell S_q}{\ell-1}\right)I_\ell\Bigg)B_\ell^T\Bigg] \\ &= \frac{1}{m^2\zeta\ell^k}\text{BlkDiag}\Bigg[B_\ell\left(\frac{2\ell S_q}{\ell-1}I_\ell - \frac{2S_q}{\ell-1}J_\ell\right)B_\ell^T, B_\ell\left(\frac{2\ell S_q}{\ell-1}I_\ell - \frac{2S_q}{\ell-1}J_\ell\right)B_\ell^T, \\ &\qquad \ldots, B_\ell\left(\frac{2\ell S_q}{\ell-1}I_\ell - \frac{2S_q}{\ell-1}J_\ell\right)B_\ell^T\Bigg].\end{aligned}$$

However, $B_\ell J_\ell B_\ell^T = \mathbf{0}$ and $B_\ell I_\ell B_\ell^T = I_{\ell-1}$, so this simplifies to

$$\sum_{\alpha=1}^{\zeta}\sum_{i<j}(C(\boldsymbol{\pi}_0)_\alpha)_{i,j} = \frac{1}{m^2\zeta\ell^k}\text{BlkDiag}\left[\frac{2\ell S_q}{\ell-1}I_{\ell-1}, \frac{2\ell S_q}{\ell-1}I_{\ell-1}, \ldots, \frac{2\ell S_q}{\ell-1}I_{\ell-1}\right].$$

Since $C(\boldsymbol{\pi}_0)_M = \sum_{\alpha=1}^{\zeta}\sum_{i<j}(C(\boldsymbol{\pi}_0)_\alpha)_{i,j}$, we have

$$\begin{aligned}C(\boldsymbol{\pi}_0)_M &= \frac{1}{m^2\zeta\ell^k}\text{BlkDiag}\left[\frac{2\ell S_q}{\ell-1}I_{\ell-1}, \frac{2\ell S_q}{\ell-1}I_{\ell-1}, \ldots, \frac{2\ell S_q}{\ell-1}I_{\ell-1}\right] \\ &= \frac{2\ell S_q}{m^2\zeta\ell^k(\ell-1)}I_{k(\ell-1)}.\end{aligned}$$

Note that when the design parameters are fixed (i.e. m, k, and ℓ) there are two variables in the above equation, S and ζ. Therefore the determinant of the information matrix is maximised when $\frac{S_q}{\zeta}$ is maximised.

If we recall that

$$\frac{S_q}{\zeta} = \frac{1}{2}\left((x+1)xy + x(x-1)(\ell-1)\right),$$

then we can see that the determinant of the information matrix is maximised when the number of non–zero differences in each generator is maximised. Equivalently, the D–optimal design will be that which minimises the number of zero differences. □

From the proof of Theorem 6.4.1 we can see that the determinant of the information matrix is

$$\det(C(\boldsymbol{\pi}_0)_M) = \left(\frac{2\ell S_q}{m^2\zeta\ell^k(\ell-1)}\right)^{k(\ell-1)}.$$

We can determine the optimal value for this determinant by finding the least upper bound for $\frac{S_q}{\zeta}$. Notice that within each set of generators, the maximum number of differences for a given m, k, and ℓ is equal to $2S_q$, and is independent of the other sets of generators. This is the case since the designs obtained by each set of generators are adjoined as extra runs of the design, thus acting independently. Then for an arbitrary ζ, Theorem 1 of Burgess and Street [2005] showed that

$$S_q = \begin{cases} \zeta(m^2-1)/4, & \ell = 2,\ m \text{ odd;} \\ \zeta m^2/4, & \ell = 2,\ m \text{ even;} \\ \zeta(m^2-(\ell x^2+2xy+y))/2, & 2 < \ell < m; \\ \zeta m(m-1)/2, & \ell \geq m. \end{cases}$$

Using this, we can find the optimal value for the determinant of the information matrix.

■ **THEOREM 6.4.2.**
The maximum value for the determinant of the information matrix when only main effects are of interest and the design is symmetric where ℓ is prime is

$$\det(C(\boldsymbol{\pi}_0)_{M,opt}) = \begin{cases} \left(\frac{m^2-1}{2^k m^2}\right)^k, & \ell = 2,\ m \text{ odd;} \\ \left(\frac{1}{2^k}\right)^k, & \ell = 2,\ m \text{ even;} \\ \left(\frac{1}{\ell^{k-1}(\ell-1)} - \frac{\ell x^2+2xy+y}{m^2\ell^{k-1}(\ell-1)}\right)^{k(\ell-1)}, & 2 < \ell < m; \\ \left(\frac{m-1}{m\ell^{k-1}(\ell-1)}\right)^{k(\ell-1)}, & \ell \geq m, \end{cases}$$

for a given m, k, and ℓ. □

Proof. Theorem 1 of Burgess and Street [2005] show that the least upper bound for $\frac{S_q}{\zeta}$ will be

$$\frac{S_q}{\zeta} = \begin{cases} (m^2-1)/4, & \ell = 2,\ m \text{ odd;} \\ m^2/4, & \ell = 2,\ m \text{ even;} \\ (m^2-(\ell x^2+2xy+y))/2, & 2 < \ell < m; \\ m(m-1)/2, & \ell \geq m. \end{cases}$$

Then if we substitute this upper bound into $\det(C(\boldsymbol{\pi}_0)_M)$, we obtain

$$\begin{aligned} \det(C(\boldsymbol{\pi}_0)_{M,\text{opt}}) &= \left(\frac{2\ell}{m^2\ell^k(\ell-1)}\left(\frac{S_q}{\zeta}\right)_{\text{opt}}\right)^{k(\ell-1)} \\ &= \begin{cases} \left(\frac{2\ell(m^2-1)/4}{m^2\ell^k(\ell-1)}\right)^{k(\ell-1)}, & \ell = 2,\ m \text{ odd;} \\ \left(\frac{2\ell m^2/4}{m^2\ell^k(\ell-1)}\right)^{k(\ell-1)}, & \ell = 2,\ m \text{ even;} \\ \left(\frac{2\ell(m^2-(\ell x^2+2xy+y))/2}{m^2\ell^k(\ell-1)}\right)^{k(\ell-1)}, & 2 < \ell < m; \\ \left(\frac{2\ell m(m-1)/2}{m^2\ell^k(\ell-1)}\right)^{k(\ell-1)}, & \ell \geq m, \end{cases} \end{aligned}$$

$$\det(C(\boldsymbol{\pi}_0)_{M,\text{opt}}) = \begin{cases} \left(\frac{m^2-1}{2^k m^2}\right)^k, & \ell = 2,\ m \text{ odd;} \\ \left(\frac{1}{2^k}\right)^k, & \ell = 2,\ m \text{ even;} \\ \left(\frac{1}{\ell^{k-1}(\ell-1)} - \frac{\ell x^2+2xy+y}{m^2\ell^{k-1}(\ell-1)}\right)^{k(\ell-1)}, & 2 < \ell < m; \\ \left(\frac{m-1}{m\ell^{k-1}(\ell-1)}\right)^{k(\ell-1)}, & \ell \geq m. \end{cases}$$

This result is the same as that obtained in Burgess and Street [2005] when a complete factorial starting design was used, and ζ was assumed to be equal to 1. □

This theorem is best illustrated in an example.

■ **EXAMPLE 6.4.3.**

In Example 6.1.1 we introduced a 3^{4-2} fractional factorial design. We can use this as a starting design for a design of a stated choice experiment. The starting design, which will form the first option in each choice set, is shown as the first column of Table 6.4.

To obtain the other options, we add generators to the starting design. In this case, let the experiment have $m = 3$ options in each choice set. Then we will need two generators to form the remaining options. According to Theorem 6.4.1, we would like to maximise the number of non–zero differences conditional on having each non–zero difference in $GF[3]$ appear equally often. The set of generators used in Example 6.3.2,

$$G = (\boldsymbol{g}_1 = (0000), \boldsymbol{g}_2 = (1111), \boldsymbol{g}_3 = (2222)),$$

yielding the design in Table 6.5, satisfy these criteria. We can confirm this by looking at each of the differences between the options. For each attribute r with $1 \leq r \leq 4$

$$\begin{array}{rclclcrclcl} g_{2,r} - g_{1,r} &=& 1-0 &=& 1 & \quad & g_{1,r} - g_{2,r} &=& 0-1 &=& 2 \\ g_{3,r} - g_{1,r} &=& 2-0 &=& 2 & & g_{1,r} - g_{3,r} &=& 0-2 &=& 1 \\ g_{3,r} - g_{2,r} &=& 2-1 &=& 1 & & g_{2,r} - g_{3,r} &=& 1-2 &=& 2. \end{array}$$

For each attribute, there are 6 non–zero differences and no zero differences. Thus $2S = 6$ and $S = 3$. Hence the information matrix for this design will be

$$\begin{aligned} C(\boldsymbol{\pi}_0)_M &= \frac{2\ell S}{m^2\ell^k(\ell-1)} I_{k(\ell-1)} \\ &= \frac{2 \times 3 \times 3}{3^2 3^4 (3-1)} I_{4\times 3} \\ &= \frac{1}{81} I_{12}, \end{aligned}$$

with determinant $\det(C(\boldsymbol{\pi}_0)_M) = (\frac{1}{81})^{12}$.

We can check whether this design is optimal by calculating $\det(C(\boldsymbol{\pi}_0)_{M,\text{opt}})$ and comparing it to the value calculated above. Using Theorem 6.4.2, the optimum value of $\det(C(\boldsymbol{\pi}_0)_M)$ for $k = 4$, $\ell = 3$, and $m = 3$ is

$$\begin{aligned} \det(C(\boldsymbol{\pi}_0)_{M,\text{opt}}) &= \left(\frac{m-1}{m\ell^{k-1}(\ell-1)}\right)^{k(\ell-1)} \\ &= \left(\frac{2}{3 \times 3^3 \times 2}\right)^{4\times 3} \\ &= \left(\frac{1}{81}\right)^{12}. \end{aligned}$$

Therefore this design is optimal for the estimation of main effects. □

Chapter 7

Conclusions and Research Directions

In this chapter, we provide a summary of our results, and discuss some possible future research directions arising from this thesis. This thesis aimed to increase the body of results on optimal designs by considering models and design approaches that provide more realistic, and less burdensome choice experiments.

In Chapter 1 we presented a summary of previously known results. Specifically we introduced the Bradley–Terry model for paired comparisons, and the MNL model for multiple comparisons. We presented results that allow researchers to find optimal designs for the estimation of main effects for both symmetric and asymmetric designs, as well as results that gave optimal designs for the estimation of main effects plus two–factor interactions for 2^k experiments when using these models. Chapter 1 also introduced paired comparison models that incorporated ties and position effects.

In Chapter 2, we derived the information matrix for the estimation of the contrasts $B_h\boldsymbol{\gamma}$ when the Davidson ties model is used. There had been no previous work on the optimal design of experiments that incorporated ties. We used this information matrix to show that the designs that are optimal for the estimation of the contrasts $B_h\boldsymbol{\gamma}$ when using the Bradley–Terry model are also optimal for the estimation of the contrasts $B_h\boldsymbol{\gamma}$ and the ties parameter when the Davidson ties model is used.

We used the equivalence result found in Chapter 2 to find optimal designs for specific sets of contrasts when the Davidson ties model is used. We gave results that allow researchers to find optimal designs for the estimation of the main effects of the attributes and ν where attributes may have any number of levels. We also established rules for finding optimal designs for the estimation of the main effects plus two–factor interactions of the attributes and ν for 2^k experiments. We used simulations to show that these designs led to parameter estimates that were unbiased and symmetrically distributed.

In Chapter 3, we introduced a generalisation of the Davidson ties model that allows an arbitrary choice set size. This generalisation was analogous to the generalisation of the Bradley–Terry model to the MNL model. In this chapter, we derived the normal equations for the maximum likelihood estimators, and the information matrix for the estimation of the contrasts in $B_h\boldsymbol{\gamma}$. Once again, we used the information matrix to show that the optimal design for the estimation of the contrasts in $B_h\boldsymbol{\gamma}$ when the MNL model is used is also optimal for the estimation of the contrasts in $B_h\boldsymbol{\gamma}$ and ν then the generalised Davidson ties model is used.

The result showing that the optimal designs for both models are the same was applied to specific sets of effects. We used the equivalence result to find rules giving the optimal designs for the estimation of the main effects of the attributes and ν where attributes may have any number of levels. We also used the equivalence result to generate rules to find the optimal design for the estimation of the main effects plus two–factor interactions of the attributes and ν for 2^k experiments. Again, we used simulations to show that these designs led to parameter estimates that were unbiased and symmetrically distributed.

One sensible extension to the work on ties in this thesis is to investigate what happens when the ties parameter depends on the number of items considered indistinguishable. That is, we may want $\nu_2, \nu_3, \ldots, \nu_m$ to reflect the fact that there are i items in the choice set that the respondent finds equally attractive, for $2 \leq i \leq m$. If this is the case, we would need to revisit the convergence criteria, as well as optimal design results.

Another research direction would be to consider different methods of incorporating ties into choice models. For example, we may wish to define a separate π parameter for finding each subset of the choice set equally attractive. Comparisons between ties models will require some sort of goodness–of–fit test to determine which approach captures preference behaviour most effectively and most parsimoniously. If we find that a different approach yields more effective models, then we will need to consider convergence criteria for the model, maximum likelihood estimators and optimal designs for the new model form.

In Chapter 4 we derived the information matrix for the estimation of the contrasts in $B_h\boldsymbol{\gamma}$ and the position main effect when the Davidson–Beaver position effects is used, allowing us to incorporate the position of the item in a choice set into the selection probability of the item. We found that the optimal designs for the estimation of the contrasts in $B_h\boldsymbol{\gamma}$ when the Bradley–Terry model is used were also optimal for the estimation of the contrasts in $B_h\boldsymbol{\gamma}$ and the position main effects when the Davidson–Beaver position effects model is used.

We used this equivalence result above to find rules to give optimal designs for the estimation of the main effects of the attributes and the position main effect where attributes may have any number of levels. We also used the result to find rules that give optimal designs for the estimation of the main effects plus two–factor interactions of the attributes and the position main effect for 2^k experiments. Once again, we used simulations to show that the parameter estimates obtained from these designs were unbiased and symmetrically distributed.

In Chapter 5 we introduced a generalisation of the Davidson–Beaver position effects model that allows for an arbitrary choice set size. We derived normal equations for the maximum likelihood estimators and found an expression for the information matrix for the estimation of the main effects of the attributes and contrasts of the position effects. We used this information matrix to prove a general result for finding optimal designs for the estimation of the main effects of the attributes and contrasts of the position effects. This optimality result was over the set of competing designs that included all choice sets characterised by a particular ordered difference vector if that difference vector was said to be included in the design. We also found an expression for the information matrix for the estimation of the main effects plus two–factor interactions of the attributes and contrasts of the position effects.

Chapter 5 also discussed designs obtained by embedding an orthogonal array into a complete

Latin square. This type of design does not belong to the class of competing designs that were considered earlier in the chapter, since the inclusion of one choice set characterised by a particular difference vector did not mean that all choice sets characterised by that difference vector are included in these designs. We found that, on some occasions, this type of design was more efficient than the optimal design from the class of designs considered earlier. We did find, however, that in order to maintain a diagonal C matrix for a 2^k design using this approach, we needed to place restrictions on the number of attributes that we could include in the design, and how they are defined. Once again, we used simulations to show that the parameter estimates obtained from these designs were unbiased and symmetrically distributed.

Further research could consider the properties of the orthogonal arrays that can be embedded into a complete Latin square to obtain a diagonal C matrix. We could also look at the properties of designs constructed in this way that have more than two levels in each attribute.

Chapter 5 did not include any results on the optimal design of choice experiments for the estimation of the main effects plus two–factor interactions of the attributes and contrasts of the position effects when the generalised Davidson–Beaver model is used. Future research could consider this problem, first for 2^k experiments, and then for a general asymmetric design. The latter of these research directions would depend on the establishment of a similar result when the MNL model is used, a result that has not been proven to date.

We could also consider variations of the position effects model. One such variation would be to consider the influence of an adjacent item in the choice set on the perceived attractiveness of the item. For example in a taste testing task, the presentation of a sample with a strong flavour could reasonably have an effect on the taste of subsequent items. We anticipate that the model form in this case will be different, as which items are adjacent to each other is now important, and not just the position in the choice set. Research into modelling this situation would involve the development of an appropriate utility function, deriving normal equations and the C matrix, and finding results that give optimal designs.

Burgess and Street [2005] have a result that uses a complete factorial as a starting design. In Chapter 6 we extended this result so that a fractional factorial design could be used as the starting design. Specifically, we consider symmetric designs with a prime power number of levels and are constructed using the Rao–Hamming method. We used permutation matrices to express the contrast matrix for the items in the first position of each choice set. Additional permutation matrices were then used to obtain the contrast matrices for the remaining options in each choice set. These contrast matrices were then used to derive an expression for the information matrix for the estimation of main effects. We then used the information matrix to prove a result that gives optimal designs for the estimation of main effects in ℓ^k experiments, where ℓ is a prime power when an orthogonal array is used as a starting design.

The results in Chapter 6 allow us to present choice experiments with fewer choice sets to the respondents, without sacrificing design efficiency. This reduces the burden placed on the respondent, and should lead to more considered and more consistent choices made by respondents due to reduced fatigue.

The designs that are considered in Chapter 6 are quite restrictive in the number of levels each attribute may take. The design optimality results only apply to symmetric experiments with a prime power number of levels. We can construct orthogonal arrays that have non-prime power numbers of levels, or are asymmetric. Future research could adapt the methods used here to extend the design optimality result to designs that use these other orthogonal arrays as a starting design.

There are constructions other than the Rao–Hamming construction that yield orthogonal arrays, linear or otherwise. For example, orthogonal arrays can be constructed from difference schemes, Hadamard matrices, and codes. Future research could look at how to find similar design optimality results when these orthogonal arrays are used as the starting design.

In this thesis, we made several assumptions such as the an independently and identically distributed Extreme Value type I error distribution. In Chapter 1 we linked this assumption to the assumption of independence from irrelevant attributes. We also presented an example where independence from irrelevant attributes was not a sensible assumption. Further research could be devoted to looking at the robustness of the optimal designs presented here when such assumptions are violated. For example, we could look at optimal designs in cases where the error distribution is correlated with the error distributions of other items in the choice set. We could also look at correlation structures between choice sets. The other aspect of this assumption that we may wish to test is the error distribution itself. We could test the effect of different error distributions, including skewed distributions, on the efficiency of the optimal designs.

There are also some more complex choice models that have not been considered in this thesis. These models go some way to accommodating the violations in assumptions mentioned earlier. Such models include the mixed logit model, the nested logit model, and models with alternative specific constants. There may be some scope to incorporate ties and position effects into these models. In addition, there is still a lot of scope for developing theory relating to optimal designs for these models.

The final research direction we consider here is models that incorporate both ties and position effects. Recall that Davidson and Beaver [1977] proposed a paired comparisons model that incorporated both of these effects. One curious feature about this model is that the utility of finding the two items in the choice set equally attractive was independent of the positions the items take. For this reason, it is not obvious how this model might be generalised for an arbitrary choice set size. Future research may be directed towards finding a model form that incorporated the respective positions of the items within the choice set to the utility of finding the items equally attractive. Such a model may have a more intuitive generalisation. Once we have these models, we could consider which designs are most efficient for the estimation of the contrasts in $B_h\boldsymbol{\gamma}$, plus ties and position effects.

Bibliography

A.C. Atkinson, A.N. Donev, and R.D. Tobias. *Optimal Experimental Designs, with SAS*. Oxford University Press, Oxford, 2007.

R.A. Bailey. *Design of Comparative Experiments*. Cambridge University Press, Cambridge, 2008.

R.J. Beaver and D.V. Gokhale. A Model to Incorporate Within–Pair Order Effects in Paired C omparisons. *Communications in Statistics*, 4(10):923–939, 1975.

K.P. Bogart. *Introductory Combinatorics*. Harcourt Academic Press, Orlando FL, 2000.

R.A. Bradley. Incomplete Block Rank Analysis: On the Appropriateness of the Model for a Method of Paired Comparisons. *Biometrics*, 10(3):375–390, Sep 1954a.

R.A. Bradley. Rank Analysis of Incomplete Block Designs: II. Additional Tables for the Method of Paired Comparisons. *Biometrika*, 41(3/4):502–537, Dec 1954b.

R.A. Bradley. Rank Analysis of Incomplete Block Designs: III. Some Large–Sample Results on Estimation and Power for a Method of Paired Comparisons. *Biometrika*, 42(3/4):450–470, Dec 1955.

R.A. Bradley and A.T. El-Helbawy. Treatment Contrasts in Paired Comparisons: Basic Procedures with Applications to Factorials. *Biometrika*, 63(2):255–262, Aug 1976.

R.A. Bradley and J.J. Gart. The Asymptotic Properties of ML Estimators when Sampling from Associated Populations. *Biometrika*, 49(1/2):205–214, Jun 1962.

R.A. Bradley and M.E. Terry. Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons. *Biometrika*, 39(3/4):324–345, Dec 1952.

L.B. Burgess and D.J. Street. Optimal Designs for 2^k Choice Experiments. *Communications in Statistics – Theory and Methods*, 32(11):2185–2206, Nov 2003.

L.B. Burgess and D.J. Street. Optimal Designs for Choice Experiments with Asymmetric Attributes. *Journal of Statistical Planning and Inference*, 134(1):288–301, Sep 2005.

L.G. Chouinard, R. Jajcay, and S.S. Magliveras. Finite Groups and Designs. In C.J. Colbourn and J.H. Dinitz, editors, *Handbook of Combinatorial Designs*, pages 819–847. CRC Press, Boca Raton FL, 2007.

K. Chrzan. Three Kinds of Order Effects in Choice-Based Conjoint Analysis. *Marketing Letters*, 5(2):165–172, 1994.

D.E. Critchlow and M.A. Fligner. Paired Comparison, Triple Comparison, and Ranking Experiments as Generalized Linear Models, and their Implementation on GLIM. *Psychometrika*, 56 (3):517–533, Sep 1991.

R.R. Davidson. On Extending the Bradley-Terry Model to Accommodate Ties in Paired Comparison Experiments. *Journal of the American Statistical Association*, 65(329):317–328, Mar 1970.

R.R. Davidson and R.J. Beaver. On Extending the Bradley-Terry Model to Incorporate Within-Pair Order Effects. *Biometrics*, 33(4):693–702, Dec 1977.

J. de Dios Ortúzar, A. Iacobelli, and C. Valeze. Estimating Demand for a Cycle-Way Network. *Transportation Research Part A*, 34:353–373, 2000.

J. Dénes and A.D. Keedwell. *Latin Squares and their Applications.* English Universities Press, London, 1974.

O. Dykstra, Jr. Rank Analysis of Incomplete Block Designs: A Method of Paired Comparisons Employing Unequal Repetitions on Pairs. *Biometrics*, 16(2):176–188, Jun 1960.

A.T. El-Helbawy. Asymptotic Relative Efficiency of Designs for Factorial Paired Comparison Experiments. *Journal of Statistical Planning and Inference*, 10(1):105–113, Jul 1984.

A.T. El-Helbawy and E.A. Ahmed. Optimal Design Results for 2^n Factorial Paired Comparison Experiments. *Communications in Statistics - Theory and Methods*, 13(22):2827–2845, 1984.

A.T. El-Helbawy and R.A. Bradley. Treatment Contrasts in Paired Comparisons: Large-Sample Results, Applications, and some Optimal Designs. *Journal of the American Statistical Association*, 73(364):831–839, Dec 1978.

A.T. El-Helbawy, E.A. Ahmed, and A.H. Alharbey. Optimal Designs for Asymmetrical Factorial Paired Comparison Experiments. *Communications in Statistics - Simulation and Computation*, 23(3):663–681, 1994.

A.B. Evans. Complete Mappings and Sequencings of Finite Groups. In C.J. Colbourn and J.H. Dinitz, editors, *Handbook of Combinatorial Designs*, pages 345–352. CRC Press, Boca Raton FL, 2007.

S.E. Fienberg. Log Linear Representation for Paired and Multiple Comparisons Models with Ties and Within-Pair Order Effects. *Biometrics*, 35(2):479–481, Jun 1979.

S.E. Fienberg and K. Larntz. Log Linear Representation for Paired and Multiple Comparisons Models. *Biometrika*, 63(2):245–254, Aug 1976.

L.R. Ford, Jr. Solution of a Ranking Problem from Binary Comparisons. *The American Mathematical Monthly*, 64(8 Part 2):28–33, Oct 1957.

W.A. Glenn and H.A. David. Ties in Paired-Comparison Experiments using a Modified Thurstone-Mosteller Model. *Biometrics*, 16(1):86–109, Mar 1960.

B. Gordon. Sequences in Groups with Distinct Partial Products. *Pacific Journal of Mathematics*, 11:1309–1313, 1961.

U. Graßhoff and R. Schwabe. Optimal Designs for the Bradley-Terry Paired Comparison Model. *Statistical Methods and Applications*, 17(3):275–289, Jul 2008.

U. Graßhoff, H. Großman, H. Holling, and R. Schwabe. Optimal paired comparison designs for first-order interactions. *Statistics*, 37(5):373–386, Oct 2003.

U. Graßhoff, H. Großman, H. Holling, and R. Schwabe. Optimal Designs for Main Effects in Linear Paired Comparison Models. *Journal of Statistical Planning and Inference*, 126(1): 361–376, Nov 2004.

U. Graßhoff, R. Schwabe, and S.G. Gilmour. Designs for first order interactions in choice experiments with binary attributes. *Otto-von-Guericke-Universität Madgeburg Facultät fur Mathematik*, 2007. URL `http://www.math.uni-magdeburg.de/~schwabe/Preprints/2007_22.pdf`.

J. P. C. Grutters, A. G. H. Kessels, C. D. Dirksen, D. van Helvoort-Postulart, L. J. C. Anteunis, and M. A. Joore. Willingness to Accept versus Willingness to Pay in a Discrete Choice Experiment. *Value in Health*, 11(7):1110–1119, 2008.

A.S. Hedayat, N.J.A. Sloane, and J. Stufken. *Orthogonal Arrays: Theory and Applications*. Springer–Verlag, New York, 1999.

R.A. Horn and C.A. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.

G. Kalton, M. Collins, and L. Brook. Experiments in Wording Opinion Questions. *Applied Statistics*, 27(2):149–161, 1978.

R.O. Kuehl. *Design of Experiments: Statistical Principles of Research Design and Analysis*. Duxbury Press, Pacific Grove CA, 2000.

W.F. Kuhfeld. *Marketing Research Methods in SAS*. SAS Institute, Cary NC, 2005.

R.C. Littell and J.M. Boyett. Designs for $R \times C$ Factorial Paired Comparison Experiments. *Biometrika*, 64(1):73–77, Apr 1977.

J. J. Louviere, D. A. Hensher, and J. D. Swait. *Stated Choice Methods: Analysis and Application*. Cambridge University Press, Cambridge, 2000.

R.D. Luce. *Individual Choice Behaviour: A Theoretical Analysis*. Wiley, New York, 1959.

I.D. Macdonald. *The Theory of Groups*. Oxford University Press, London, 1968.

D. McFadden. Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka, editor, *Frontiers in Econometrics*, pages 105–142. Academic Press, New York, 1973.

K.A. Phillips, T. Maddala, and F. Reed-Johnson. Measuring Preferences for Health Care Interventions using Conjoint Analysis: An Application to HIV Testing. *Health Services Research*, 37(6):1681–1705, Dec 2002.

M.H. Quenouille and J.A. John. Paired Comparison Designs for 2^n Factorials. *Applied Statistics*, 20(1):16–24, 1971.

C.R. Rao. Factorial Experiments Derivable from Combinatorial Arrangements of Arrays. *Journal of the Royal Statistical Society (Supp)*, 9:128–139, 1947.

C.R. Rao. On a Class of Arrangements. *Proceesings of the Edinburugh Mathematical Society*, 8: 119–125, 1949.

P.V. Rao and L.L. Kupper. Ties in Paired–Comparison Experiments: A Generalization of the Bradley-Terry model. *Journal of the American Statistical Association*, 62(317):194–204, Mar 1967.

A.P. Street and D.J. Street. *Combinatorics of Experimental Design.* Oxford University Press, New York, 1986.

D.J. Street. Factorial Designs. In C.J. Colbourn and J.H. Dinitz, editors, *Handbook of Combinatorial Designs*, pages 445–465. CRC Press, Boca Raton FL, 2007.

D.J. Street and L. Burgess. *The Construction of Optimal Stated Choice Experiments: Theory and Methods.* John Wiley & Sons, Hoboken NJ, 2007.

D.J. Street, D.S. Bunch, and B.J. Moore. Optimal Designs for 2^k Paired Comparison Experiments. *Communications in Statistics – Theory and Methods*, 30(10):2149–2171, Oct 2001.

D.J. Street, L.B. Burgess, and J.J. Louviere. Quick and Easy Choice Sets: Constructing Optimal and Nearly Optimal Stated Choice Experiments. *International Journal of Research in Marketing*, 22(4):459–470, Dec 2005.

M. Tharp and L. Marks. An Examination of the Effects of Attribute Order and Product Order Biases in Conjoint Analysis. *Advances in Consumer Research*, 17:563–570, 1990.

L. L. Thurstone. The Method of Paired Comparisons for Social Values. *Journal of Abnormal and Social Psychology*, 21:384–400, 1927.

K.E. Train. *Discrete Choice Methods with Simulation.* Cambridge University Press, New York, 2003.

E.E.M van Berkum. *Optimal Paired Comparison designs for factorial experiments.* CWI Tract, 1987.

P. van der Waerden, A. Borgers, H. Timmermans, and M. Bérénos. Order Effects in Stated-Choice Experiments: Study of Transport Mode Choice Decisions. *Transportation Research Record*, 1985:12–18, 2006.

E.J. Williams. Experimental Designs Balanced for the Estimation of Residual Effects of Treatments. *Australian Journal of Scientific Research Series A*, 2:149–168, 1949.

E. Zermelo. Die Berechnung der Turnier–Ergebnisse als ein Maximumproblem der Wahrschreinlichkeitsrechnung. *Mathemalische Zeitchrift*, 29:436–460, 1929.