Relative efficiencies of alternative designs for randomized trials

RELATIVE EFFICIENCIES OF ALTERNATIVE DESIGNS FOR RANDOMIZED TRIALS

BY

MENGJIE BIAN, B.Sc.

A THESIS

SUBMITTED TO THE DEPARTMENT OF MATHEMATICS & STATISTICS

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

Master of Science (2019)   McMaster University

(Mathematics & Statistics)   Hamilton, Ontario, Canada

TITLE: Relative efficiencies of alternative designs for randomized trials

AUTHOR: Mengjie Bian

                            B.Sc., (Economics)

                            China Jiliang University, Hangzhou, China

SUPERVISOR: Dr. Stephen Walter

NUMBER OF PAGES: x, 74

To my family

Abstract

In conventional randomized controlled trials (RCT), one cannot estimate the influences of patient preferences on the treatment outcomes since all patients are randomized to a treatment without identifying their preferences. The two-stage design allows patients in the choice group to choose their desired treatment while those in the random group are randomized in the same way as in a RCT. The partially randomized design allows all patients to receive their preferred treatment. In the fully randomized design, although the patient preferences are identified at the first stage, all the patients are then randomized to each treatment. In this thesis, we discuss these four designs in detail with respect to their estimable effects, variances of the estimable effects, and the relative efficiency of each effect in different designs. Participants who are indifferent to the treatments (undecided participants) are included in the designs when evaluating the various effects. This thesis also shows the relationships of relative efficiency to other factors, such as the proportion of undecided participants, the relative preference rate between the two treatments, and with unequal numbers of participants being randomized to each treatment. We discuss the advantages and disadvantages of the designs under different scenarios and whether unequal randomization could improve efficiency.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my advisor Prof. Walter for the continuous support of my master study and related research, for his patient, motivation, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Shui Feng and Prof. Lehana Thabane for offering valuable advice to me.

Special thanks to my friend: Wengxin, Guo for his help.

Finally, I would like to thank to my parents for their love.

Notation and abbreviations

TE...Treatment effect

SE...Selection effect

PE...Preference effect

SD...Standard deviation

SEs...Standard errors

Contents

List of Figures

Chapter 1

Introduction

The randomized controlled trial, which is also called the conventional parallel group randomized design in this thesis, is typically used to evaluate the direct treatment effect. In practice, however, the outcomes might be influenced by patients' preferences (if they were allowed to choose). Selection effect measures the difference in treatment outcomes between the participants with a different preference. The outcomes may also be affected by the interaction between the patients choice and the treatment received, which is termed as preference effect (Rücker, 1989). In clinical trials, there might also exist patients who have no preference. Walter *et al.* (2017b) proposed the so-called second selection and preference contrasts, which measure the difference in outcomes between the patients with a definite preference and the patients without a preference.

Rücker (1989) proposed a two-stage randomized model, which can evaluate the selection and preference effects, which are confounded by the treatment effect in the traditional randomized trial. In this design, participants are initially allocated to one

of two groups, which are the random group and the choice group. Participants in the choice group are allowed to choose their preferred treatment and patients having no preference are randomized to one of the treatments. Participants in the random group are randomized to treatment A or treatment B. Overall, compared with the conventional parallel randomized design, the two-stage randomized design allows at least a subgroup of participants to have a free choice of treatments. The partially randomized design, which was proposed by Brewin and Bradley (1989), allows all patients to receive their desired treatments and randomizes participants with no preference. The fully randomized design was developed by Torgerson *et al.* (1996). This design is a randomized control trial which includes preference information, with all treatment preferences stated before randomization.

In this thesis, we will follow the model proposed by Rücker (1989). It is constructed as a linear combination of selection effect, preference effect and treatment effect. Few papers discussed the scenarios with undecided participants. Although the undecided participants are considered in this overall model, Rücker (1989) assumed either there are no undecided participants, or that the undecided participants are not subject to selection or preference effects. Rücker (1989) also analyzed the various effects by hypothesis testing. However, in this thesis, we are interested in the estimation of the various effects.

Walter *et al.* (2017b) extended the analysis of two-stage randomized design by considering the undecided participants. In his paper, he derived the corresponding estimates and variances of selection effect, preference effect and second contrasts for

selection and preference effects. He also showed that testable assumptions are possible to have unbiased estimation.

The two-stage design proposed by Rücker (1989) is restricted to continuous response variables. Cameron *et al.* (2018) extended the two-stage design which allows for binary outcome measures. They also extended the binary outcomes methods to include stratification factors, which may be useful to more closely mirror the distribution of preferences in the study population and may have impact on overall study efficiency.

Based on Walter 's study on the undecided participants, this thesis will extend the earlier studies to allow for undecided participants in the fully randomized design. Comparing the relative efficiencies of various designs, we aim to find an optimal design depending on the proportion of undecided participants and preference rates for the two treatments.

In the R*ü*cker's study, equal numbers of participants are randomized to each treatment. In this thesis, we will introduce a parameter ρ related to the allocation rate between two treatments to compare the differences in efficiency between unequal and unequal randomization. In practice, clinicians may be more be more familiar with one treatment than another or patients in the choice group show a greater preference to one treatment. Therefore, unequal allocation might be favored in the ethical aspect. Also, we expect unequal allocation could improve the estimation of one or more parameters of interest. For instance, when a greater portion of participants prefer one

treatment, investigators may consider allocating more participants to that treatment than the other one or instead, assigning more participants to the less preferred treatment to achieve a balance between the treatments. We want to see if this unequal allocation would enhance the efficiencies of estimations.

In the next chapter, we will introduce details of various designs, which include conventional parallel randomized design, two-stage randomized design, partially randomized design and fully randomized design. The corresponding estimable effects and their variances will also be derived. Chapter 3 will show graphs of relative efficiencies related to various effects for each design and we will also compare relative efficiencies of the effects among designs, under both equal and unequal randomization. Relative inverted standard errors will be used as a measure of relative efficiency, with a base case where undecided rate is 50% and participants have an equal preference for the two treatments. Relative efficiency varies with relative preference rates for treatments A and B and the undecided rate, so that we can evaluate the patterns of relative efficiencies under both parameters. At the end of this chapter, we will discuss whether the unequal randomization to some extent improve the estimations. Chapter 4 will illustrate an example by using Cooper's data regarding to heavy menstrual bleeding. (Cooper *et al.*, 1997, 1999) Relative preference rates for treatments and undecided rate with respect to this example will be calculated to substitute into the estimates for various designs. We will discuss whether the efficiencies of estimates could be improved by changing the undecided rates and also consider the advantages and disadvantages for various designs. In the final chapter, we will mainly discuss about the designs which are recommended under different scenarios and the advantages and

disadvantages of various designs.

Chapter 2

Methods

This chapter will first introduce a linear model and then go into the detail for the different designs, which are conventional parallel group randomized design, two stage randomized design, partially randomized design and the fully randomized design. For each design, we are interested in the impacts of three effects, which are the treatment effect (TE), selection effect (SE) and preference effect (PE). And we will also deal with the second selection and preference contrasts for the different designs. For each effect, we also consider whether it is estimable or not. At the end of this chapter, we will derive the estimated variances for different designs. Some of the deviations are based on Walter's paper (Walter $et~al.$, 2017b).

Proposed by Rücker (1989), the linear model we adopt is as follows:

$$Y_{ijk} = \mu + \tau_i + \nu_j + \pi_{ij} + \varepsilon_{ijk}$$

where Y_{ijk} represents the effect of treatment i (i=1 for treatment A, 2 for treatment

B) actually received by participant k who prefer treatment j (j=1 for treatment A, 2 for B, and j=3 for no preference), μ is the overall mean response, τ represents the treatment effect, ν stands for the selection effect, π for the preference effect, and ε is the random error term, which is assumed to be independent of other terms and have zero expectation.

Furthermore, there are four constraints required to avoid redundancy:

$$\tau_1 + \tau_2 = 0,$$

$$\alpha\nu_1 + \beta\nu_2 + \gamma\nu_3 = 0,$$

$$\alpha\pi_{i1} + \beta\pi_{i2} + \gamma\pi_{i3} = 0, \quad (i = 1, 2),$$

$$\pi_{1j} + \pi_{2j} = 0, \quad (j = 1, 2, 3)$$

where α and β are the preference rates for the treatment A and treatment B, respectively, and γ is the undecided rate (the percentage of the participants who are indifferent between the treatment A and treatment B).

The treatment effect ($\Delta\tau$) is the direct effect of treatment. It is defined as $\Delta\tau = \tau_1 - \tau_2$, or equivalently $\Delta\mu = \mu_1 - \mu_2$. The selection effect ($\Delta\nu$) measures to what extent the difference in the expected mean response is influenced by choosing one treatment or the other. It is given by $\Delta\nu = \nu_1 - \nu_2$, or equivalently, $\Delta\nu = [(\mu_{11} + \mu_{21}) - (\mu_{12} + \mu_{22})]/2$ in terms of mean responses. The preference effect, denoted by $\Delta\pi$, refers to the difference caused by the interaction between the actual treatment received and the treatment preferred by the participant. It is defined as

7

$\Delta\pi = [(\pi_{11} + \pi_{22}) - (\pi_{12} + \pi_{21})]/2$. Or equivalently, $\Delta\pi = [(\mu_{11} - \mu_{12}) - (\mu_{21} - \mu_{22})]/2$, which represents the difference in treatment effects A and B between those who receive the treatment they prefer and those who do not.

First contrasts, which are selection effect and preference effect, concern comparisons between those participants with a strong preference for a treatment. In practice, there is a portion of patients who are indifferent to treatment choice. Effects which describe comparisons between those have a preference and those without a preference are called second contrasts, which were proposed by Walter *et al.* (2017b). The second selection contrast is defined as:

$$\Delta\nu' = \nu_3 - \frac{(\nu_1 + \nu_2)}{2}$$

It measures the difference between the expected mean responses for the undecided participants and those who have a specific treatment preference.

The second preference contrast is defined as a similar way:

$$\Delta\pi' = [(\frac{(\pi_{11} + \pi_{12})}{2} - \pi_{13}) - (\frac{(\pi_{21} + \pi_{22})}{2} - \pi_{23})]/2$$

It represents the difference in treatment A and treatment B between the participants who have a preference and those with no preference.

2.1   Conventional Parallel Group Randomized Trial

Conventional parallel group randomized trial is widely accepted as a method to evaluate the effectiveness of new treatments. In this design, all participants are randomized to one of the two treatment groups. Such trial reduces some type of systematic error that may interfere with treatment effects. However, such randomization ignores participant preferences, which may affect the outcomes. (Figure 2.1)

We have observable outcomes in two groups, they are (equation (1)):
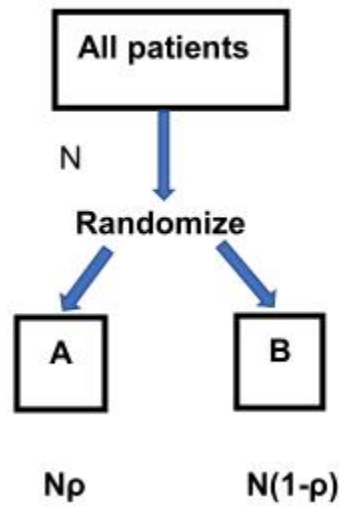
Figure 2.1: Conventional randomized trial

$$\text{Randomised to A:} \quad \mu_1 = \mu + \tau_1 + \alpha(\nu_1 + \pi_{11}) + \beta(\nu_2 + \pi_{12}) + \gamma(\nu_3 + \pi_{13})$$

$$\text{Randomised to B:} \quad \mu_2 = \mu + \tau_2 + \alpha(\nu_1 + \pi_{21}) + \beta(\nu_2 + \pi_{22}) + \gamma(\nu_3 + \pi_{23})$$

2.1.1 Estimable Effect

We can get the TE directly from the equation (1):

$$E(\Delta\hat{\tau}) = \tau_1 - \tau_2 + \alpha(\pi_{11} - \pi_{21}) + \beta(\pi_{12} - \pi_{22}) + \gamma(\pi_{13} - \pi_{23})$$

From the constraints, we see that the last three terms cancel and thus $\Delta\hat{\nu}$ is unbiased. Since the preferences for the treatment are not assessible in this design, SE and PE can not be estimated. Also, we cannot estimate the second selection and preference contrasts.

2.1.2 Variances of the Estimated TE

We assume the total sample size is N. If equal numbers of participants are allocated to treatment A and treatment B (equal randomization), then we have N/2 participants in each group. On the other hand, if unequal numbers of participants are allocated to each treatment (unequal randomization), the sample sizes for two groups are $N\rho$ and $N(1-\rho)$, where ρ represents the randomisation rate for treatment A). The variance of estimated TE is straightforward to obtain. If we assume the standard deviation of the random groups A and B (σ_1 and σ_2) are equal , then $var(\Delta\hat{\tau}) = \sigma_1^2/(N/2) + \sigma_2^2/(N/2) = 4\sigma^2/N$ in the case of equal randomization.

On the other hand, if we allow unequal randomization, then $var(\Delta\hat{\tau}) = \sigma_1^2/(N\rho) + \sigma_2^2/(N(1-\rho)) = \sigma^2/N\rho(1-\rho)$ (equation (2))

2.2   Two-stage Randomized Design

In the two-stage randomized design, participants are firstly randomized into choice group or random group. In the choice group, the participants who have a treatment preference are allowed to receive the treatment they prefer while the participants with no preference are randomized into either treatment. The participants in the random group are randomized into one treatment as the standard parallel design. (Figure 2.2)
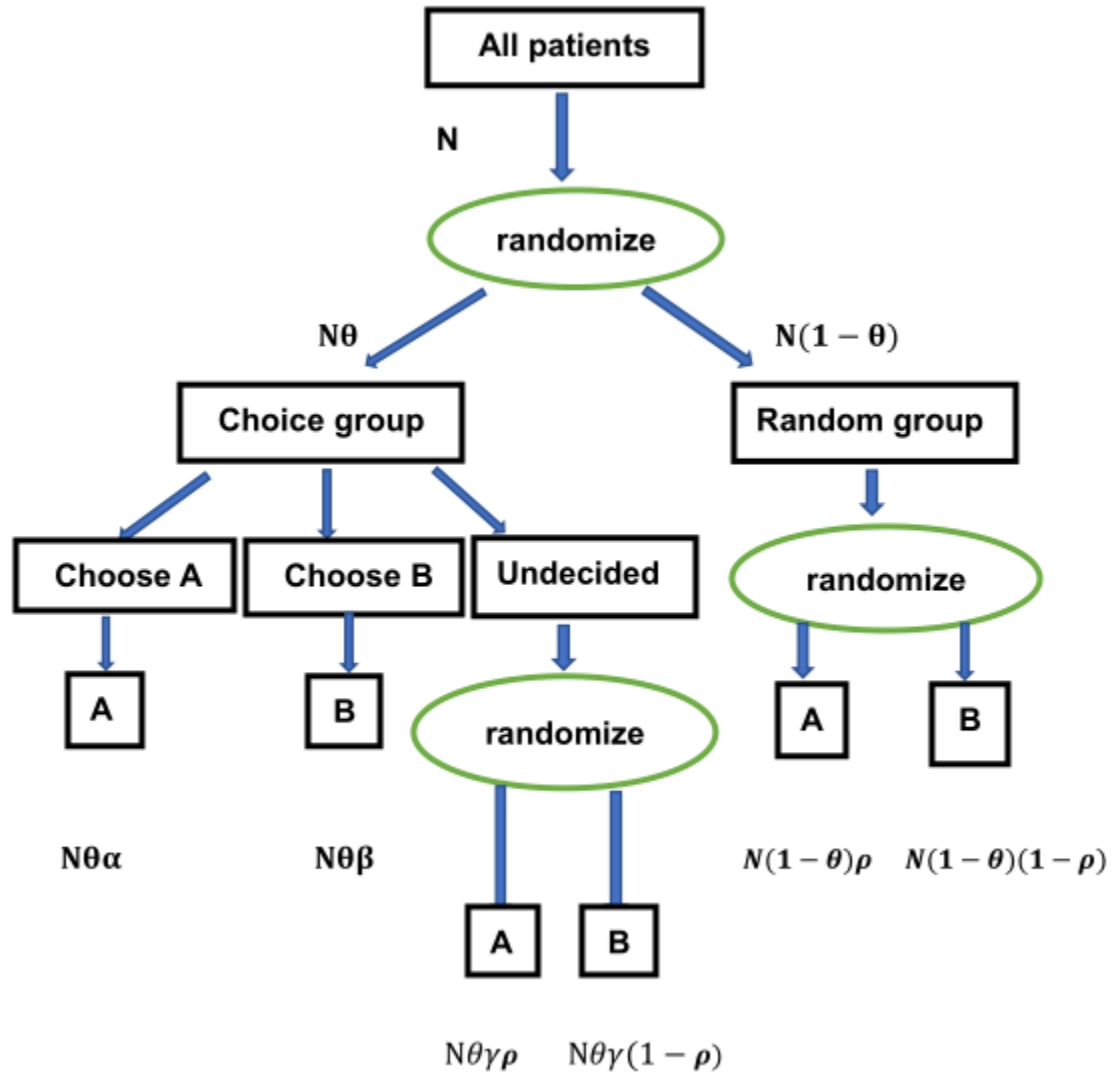
Figure 2.2: Two-stage randomised design

Then, we can get six observable groups, the expected mean responses are:

Choice arm:

$$\text{Choose A:} \qquad \mu_{11} = \mu + \tau_1 + \nu_1 + \pi_{11} \qquad (3)$$

$$\text{Choose B:} \qquad \mu_{22} = \mu + \tau_2 + \nu_2 + \pi_{22} \qquad (4)$$

$$\text{randomized to A:} \qquad \mu_{13} = \mu + \tau_1 + \nu_3 + \pi_{13} \qquad (5)$$

$$\text{randomized to B:} \qquad \mu_{23} = \mu + \tau_2 + \nu_3 + \pi_{23} \qquad (6)$$

Random arm:

$$\text{Randomised to A:} \qquad \mu_1 = \mu + \tau_1 + \alpha(\nu_1 + \pi_{11}) + \beta(\nu_2 + \pi_{12}) + \gamma(\nu_3 + \pi_{13})$$

$$\text{Randomised to B:} \qquad \mu_2 = \mu + \tau_2 + \alpha(\nu_1 + \pi_{21}) + \beta(\nu_2 + \pi_{22}) + \gamma(\nu_3 + \pi_{23})$$

2.2.1   Estimable effects

There are two estimators of TE, one is from difference between μ_1 and μ_2, the other one is available from the undecided group, which is given by:

$$E(\Delta\hat{\tau}) = \mu_{13} - \mu_{23} = \tau_1 - \tau_2 + \pi_{13} - \pi_{23}$$

$$= TE + \pi_{13} - \pi_{23} \qquad (7)$$

The first estimator $E(\Delta\hat{\tau}) = \mu_1 - \mu_2$ is unbiased, while the second estimator from the undecided group is biased by π_{ij} terms.

Walter *et al.* (2017a) gave the expressions for the estimators for the SE and PE, which are all unbiased. To derive the estimators of SE and PE, we need to get the

estimators of μ_{12} and μ_{21}, which are available from:

$$\mu_1 = \alpha\mu_{11} + \beta\mu_{12} + \gamma\mu_{13} \tag{8}$$

$$\mu_2 = \alpha\mu_{21} + \beta\mu_{22} + \gamma\mu_{23} \tag{9}$$

Since μ_{11}, μ_{22}, μ_{13}, μ_{23} are observable, by substituting for the two unobservable means in the $\Delta\nu = [(\mu_{11}+\mu_{21})-(\mu_{12}+\mu_{22})]/2$, and $\Delta\pi = [(\pi_{11}+\pi_{22})-(\pi_{12}-\pi_{21})]/2$, we could obtain :

$$\Delta\hat{\nu} = [(z_1 - z_2) - \hat{\gamma}(w_1 - w_2)]/(2\hat{\alpha}\hat{\beta}m) \tag{10}$$

$$\Delta\hat{\pi} = [(z_1 + z_2) - \hat{\gamma}(w_1 + w_2)]/(2\hat{\alpha}\hat{\beta}m) \tag{11}$$

where $z_1 = m_1(\hat{\mu_{11}} - \hat{\mu_1})$, $z_2 = m_2(\hat{\mu_{22}} - \hat{\mu_2})$, $w_1 = m_1(\hat{\mu_{11}} - \hat{\mu_{13}})$, and $w_2 = m_2(\hat{\mu_{22}} - \hat{\mu_{23}})$. (m is the sample size of choice group, m_1 is the sample size of subgroup A in which patients prefer treatment A and receive treatment A and m_2 is the sample size of subgroup B in which patients prefer treatment B and receive treatment B).

For the second contrast of selection effect, we can also define in terms of the mean responses, which is as follows:

$$\Delta\hat{\nu}' = [(\mu_{13} + \mu_{23}) - \frac{1}{2}(\mu_{11} + \mu_{21} + \mu_{12} + \mu_{22})]/2$$

In terms of estimable quantities, Walter *et al.* (2017b) derived the expression as:

$$\Delta\hat{\nu}' = [(z_1 + z_2) - (w_1 + w_2) + (\hat{\alpha} - \hat{\beta})(w_1 - w_2)]/(4\hat{\alpha}\hat{\beta}m) \tag{12}$$

The corresponding second contrast of preference effect is

$$\Delta\hat{\pi}' = [-(z_1 - z_2) + (w_1 - w_2) - (\hat{\alpha} - \hat{\beta})(w_1 + w_2)]/(4\hat{\alpha}\hat{\beta}m) \tag{13}$$

2.2.2　Variances of the Effects

Assuming θ is the proportion of patients randomized to the choice group, the number of participants randomized to choice group is $N\theta$, and N(1-θ) for the random group. In the two-stage design, we obtain the unbiased estimator of TE from the randomized group. Again, if we assume a constant variance in each group (σ^2), then the variance of the $\Delta\hat{\tau}$ is:

$$Var(\Delta\hat{\tau}) = \frac{\sigma^2}{N(1 - \theta)\rho} + \frac{\sigma^2}{N(1 - \theta)(1 - \rho)} \tag{14}$$

If we assume equal randomization, then

$$Var(\Delta\hat{\tau}) = \frac{4\sigma^2}{N(1 - \theta)} \tag{15}$$

Walter *et al.* (2017b) derived the variance of the estimators of SE and PE both unconditionally and conditionally. We regard the variances as unconditional when we take the sample variations of alpha, beta, gamma parameters into consideration. With the conditional approach, we ignore the sample variations, which means the parameters such as α, β, γ and m, m_1, m_2 are taken as taken as fixed. Since it

has been noted by Walter *et al.* (2017b) that the conditional and unconditional have similar numerical results, therefore we only consider the conditional variances in this thesis.

Thus, it shows that the estimated variance for the selection and preference effects are equal:

$$var(\Delta\hat{\nu}) = var(\Delta\hat{\pi}) = \frac{\sigma^2}{4\alpha^2\beta^2m}[(1-\gamma)^3 + \alpha^2\frac{1}{\rho}(\gamma + \frac{\theta}{1-\theta}) + \beta^2\frac{1}{1-\rho}(\gamma + \frac{\theta}{1-\theta})] \quad (16)$$

Again, if it is equal randomization, then

$$var(\Delta\hat{\nu}) = var(\Delta\hat{\pi}) = \frac{\sigma^2}{4\alpha^2\beta^2m}[(1-\gamma)^3 + 2(\alpha^2+\beta^2)(\gamma + \frac{\theta}{1-\theta})]$$

The estimated variances for the second selection and preference contrasts are:

$$var(\Delta\hat{v}') = var(\Delta\hat{\pi}') = \frac{\sigma^2}{16\alpha^2\beta^2\gamma m}[\gamma(1-\gamma)(\alpha-\beta)^2 + \frac{\alpha^2}{\rho}(2\beta+\gamma)^2 + \frac{\beta^2}{1-\rho}(2\alpha+\gamma)^2$$
$$+ \gamma(\frac{\alpha^2}{\rho} + \frac{\beta^2}{1-\rho})(\frac{\theta}{1-\theta})] \quad (17)$$

In the case of equal randomization (Walter *et al.*, 2017b), it would be:

$$var(\Delta\hat{v}') = var(\Delta\hat{\pi}') = \frac{\sigma^2}{16\alpha^2\beta^2\gamma m}[\gamma(1-\gamma)(\alpha-\beta)^2 + 2\{\alpha^2(2\beta+\gamma)^2$$
$$+ \beta^2(2\alpha+\gamma)^2\} + 2\gamma(\alpha^2+\beta^2)(\frac{\theta}{1-\theta})]$$

2.3   Partially Randomized Preference Design

The process of partially randomized preference design is shown in Figure 2.3. In this design, all participants are asked whether they have a preference. If they do have a preference, then they will receive the treatment they prefer. Those who have no specific preference are randomized to one of the two treatments.
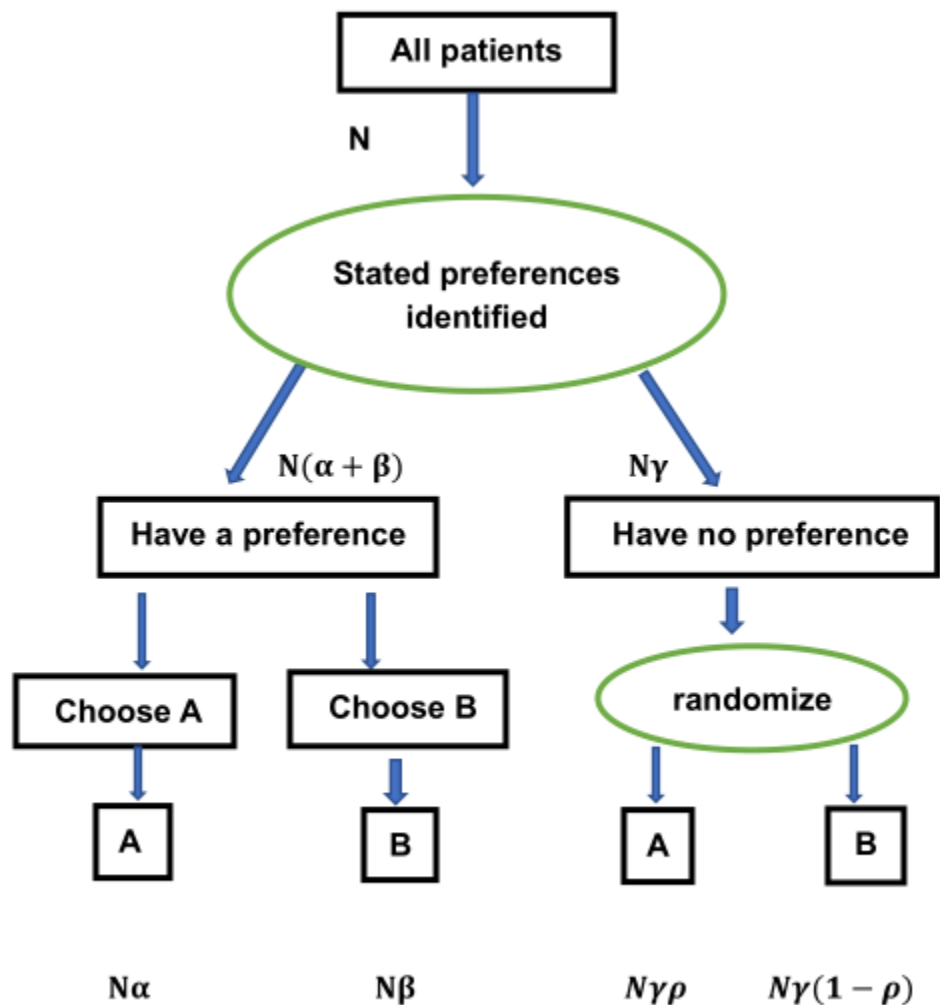
Figure 2.3: Partially randomised design

As shown in the figure 2.3, we have four observable groups, which have the expected mean responses as follows:

Prefer A, receive A:　　　　　　　　　$\mu_{11} = \mu + \tau_1 + \nu_1 + \pi_{11}$

Prefer B, receive B:　　　　　　　　　$\mu_{22} = \mu + \tau_2 + \nu_2 + \pi_{22}$

Undecided, randomized to A:　　　　　$\mu_{13} = \mu + \tau_1 + \nu_3 + \pi_{13}$

Undecided, randomized to B:　　　　　$\mu_{23} = \mu + \tau_2 + \nu_3 + \pi_{23}$

2.3.1　Estimable Effects

The estimator of TE is available through the undecided group:

$$E(\Delta\hat{\tau}) = \mu_{13} - \mu_{23} = \tau_1 - \tau_2 + (\pi_{13} - \pi_{23}) \tag{18}$$

So, we can see that this estimator is biased unless $\pi_{13} = \pi_{23}$.

The possible estimator for the SE could be obtained from:

$$E(\Delta\hat{\nu}) = (\mu_{11} - \mu_{13}) - (\mu_{22} - \mu_{23}) = \nu_1 - \nu_2 + (\pi_{11} - \pi_{13} - \pi_{22} + \pi_{23}) \tag{19}$$

This estimator is also biased if $\pi_{11} - \pi_{13} - \pi_{22} + \pi_{23}$ is not equal to zero. Otherwise, if $\pi_{11} = \pi_{13}$ and $\pi_{22} = \pi_{23}$, then it means that there is no difference in the mean responses to each treatment between those who prefer it and who are randomized to it. (Walter *et al.*, 2017a)

PE cannot be estimated in this design since π_{12} and π_{21} are not are not available from the equation (16).

The second selection contrast could be estimated from:

$$
\begin{aligned}
E(\Delta\hat{\nu}') &= \frac{\mu_{13} + \mu_{23}}{2} - \frac{\mu_{11} + \mu_{22}}{2} \\
&= \nu_3 - \frac{\nu_1 + \nu_2}{2} + \frac{1}{2}[(\pi_{13} + \pi_{23}) - (\pi_{11} + \pi_{22})]
\end{aligned}
\tag{20}
$$

which is unbiased only when $(\pi_{13} + \pi_{23}) - (\pi_{11} + \pi_{22}) = 0$. If it turns out to be zero, then it means that there is no difference in the expected mean responses to each treatment between those who prefer it and who are randomized to it.

2.3.2 Variances of the Effects

We estimate the TE from the undecided group, with size of $N\gamma\rho$ for the treatment A and $N\gamma(1-\rho)$ for the treatment B. Thus, the variance of the estimated TE is:

$$
var(\Delta\hat{\tau}) = \frac{\sigma^2}{N\gamma\rho} + \frac{\sigma^2}{N\gamma(1-\rho)} = \frac{\sigma^2}{N\gamma\rho(1-\rho)}
\tag{21}
$$

The estimated SE involves terms of estimated μ_{11}, μ_{13}, μ_{22} and μ_{23}. They have the corresponding sample sizes of $N\alpha$, $N\gamma\rho$, $N\beta$, and $N\gamma(1-\rho)$. Thus, the estimated selection effect leads to:

$$var(\Delta\hat{\nu}) = var[(\hat{\mu}_{11} - \hat{\mu}_{13}) - (\hat{\mu}_{22} - \hat{\mu}_{23})]$$

$$= \sigma^2(\frac{1}{N\alpha} + \frac{1}{N\gamma\rho} + \frac{1}{N\beta} + \frac{1}{N\gamma(1-\rho)})$$

$$= \sigma^2(\frac{1}{N\alpha} + \frac{1}{N\gamma}\frac{1}{\rho(1-\rho)} + \frac{1}{N\beta})$$

$$= \frac{\sigma^2}{N}(\frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma}\frac{1}{\rho(1-\rho)}) \tag{22}$$

If we use equal randomization, then

$$var(\Delta\hat{\nu}) = \sigma^2(\frac{1}{N\alpha} + \frac{1}{\frac{N\gamma}{2}} + \frac{1}{N\beta} + \frac{1}{\frac{N\gamma}{2}})$$

$$= \sigma^2(\frac{1}{N\alpha} + \frac{4}{N\gamma} + \frac{1}{N\beta})$$

$$= \frac{\sigma^2}{N}(\frac{1}{\alpha} + \frac{1}{\beta} + \frac{4}{\gamma})$$

Since the estimated PE is not available, thus the variance of it cannot be estimated in this design.

For the estimated variance of second selection contrast, we could obtain from:

$$var(\Delta\hat{\nu}') = var(\frac{\hat{\mu}_{13} + \hat{\mu}_{23}}{2} - \frac{\hat{\mu}_{11} + \hat{\mu}_{22}}{2})$$

$$= \frac{1}{4}var(\hat{\mu}_{13} + \hat{\mu}_{23}) + \frac{1}{4}var(\hat{\mu}_{11} + \hat{\mu}_{22})$$

$$= \frac{1}{4}\sigma^2(\frac{1}{N\gamma\rho} + \frac{1}{N\gamma}\frac{1}{(1-\rho)} + \frac{1}{N\alpha} + \frac{1}{N\beta})$$

$$= \frac{\sigma^2}{4N}(\frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma(1-\rho)\rho}) \tag{23}$$

If we use equal randomization, then estimated variance becomes:

$$var(\Delta\hat{\nu}') = \frac{\sigma^2}{N}\left(\frac{1}{4\alpha} + \frac{1}{4\beta} + \frac{1}{\gamma}\right)$$

2.4   Fully Randomized Preference Design

The preferences of all participants are identified at first and then they are all randomized into each treatment. (Figure 2.4) In this design, we have six observable groups:

Prefer A, randomized to A: $\mu_{11} = \mu + \tau_1 + \nu_1 + \pi_{11}$

Prefer B, randomized to A: $\mu_{12} = \mu + \tau_1 + \nu_2 + \pi_{12}$

Undecided, randomized to A: $\mu_{13} = \mu + \tau_1 + \nu_3 + \pi_{13}$

Prefer A, randomized to B: $\mu_{21} = \mu + \tau_2 + \nu_1 + \pi_{21}$

Prefer B, randomized to B: $\mu_{22} = \mu + \tau_2 + \nu_2 + \pi_{22}$

Undecided, randomized to B: $\mu_{23} = \mu + \tau_2 + \nu_3 + \pi_{23}$
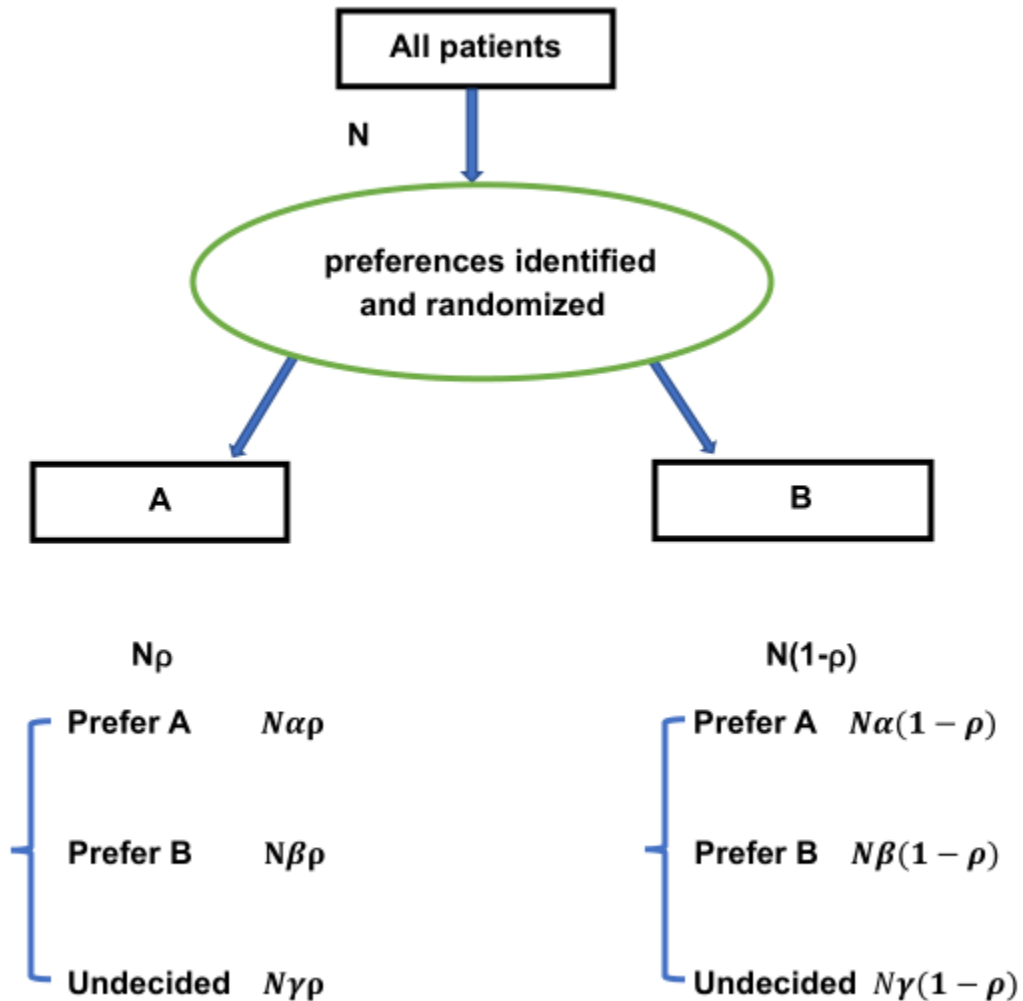
Figure 2.4: Fully randomised design

2.4.1   Estimable Effects

There are three available estimators of TE, one is from the group which prefer A, one is from the group which prefer B, the other one is from the undecided group. They

are given as follows:

$$E(\Delta\hat{\tau}_A) = E(\hat{\mu}_{11} - \hat{\mu}_{21}) = \mu_{11} - \mu_{21} = \tau_1 - \tau_2 + (\pi_{11} - \pi_{21}) \tag{24}$$

$$E(\Delta\hat{\tau}_B) = E(\hat{\mu}_{12} - \hat{\mu}_{22}) = \mu_{12} - \mu_{22} = \tau_1 - \tau_2 + (\pi_{12} - \pi_{22}) \tag{25}$$

$$E(\Delta\hat{\tau}_U) = E(\hat{\mu}_{13} - \hat{\mu}_{23}) = \mu_{13} - \mu_{23} = \tau_1 - \tau_2 + (\pi_{13} - \pi_{23}) \tag{26}$$

These three estimators are all biased by π_{ij} terms. We could also obtain an another estimator by taking the weighted averge of the above three estimators, with the weights α, β and γ, which represent the preference rate for the treatment A, treatment B, and the undecided rate, repectively. This is given by :

$$E(\alpha\Delta\hat{\tau}_A + \beta\Delta\hat{\tau}_B + \gamma\Delta\hat{\tau}_U) = \tau_1 - \tau_2 + \alpha(\pi_{11} - \pi_{21})$$
$$+ \beta(\pi_{12} - \pi_{22}) + \gamma(\pi_{13} - \pi_{23}) \tag{27}$$

According the contraints, the last three terms add up to be zero, thus this estimator is unbiased. Note that if we ignore the baseline information of preferences, then we can get same outcomes as the conventional parallel group randomised design. Thus, comparison within the subgroups (A, B, and undecided group) in the fully randomised design reveals the information which is confounded by the conventional parallel group randomised design. As shown in equation (24), (25) and (26), the treatment effect in each subgroup is biased by preference effect.

The selection effect and preference effect are also estimable. The estimator of

preference effect can be directly obtained from:

$$E(\Delta\hat{\pi}) = E[(\hat{\mu}_{11} + \hat{\mu}_{22}) - (\hat{\mu}_{12} + \hat{\mu}_{21})]/2 = [(\pi_{11} + \pi_{22}) - (\pi_{12} + \pi_{21})]/2$$

which is unbiased. The selection effect can be estimated from:

$$E(\Delta\hat{\nu}) = \frac{E[(\hat{\mu}_{11} + \hat{\mu}_{21}) - (\hat{\mu}_{12} + \hat{\mu}_{22})]}{2}$$

$$= \nu_1 - \nu_2 + \frac{1}{2}(\pi_{11} + \pi_{21} - \pi_{12} - \pi_{22})$$

which is unbiased only if $\pi_{11} + \pi_{21} - \pi_{12} - \pi_{22} = 0$. We can estimate the second selection contrast through:

$$\Delta\hat{\nu}' = [(\hat{\mu}_{13} + \hat{\mu}_{23}) - \frac{1}{2}(\hat{\mu}_{11} + \hat{\mu}_{21} + \hat{\mu}_{12} + \hat{\mu}_{22})]/2$$

Similarly, the estimator of second preference contrast is :

$$\Delta\hat{\pi}' = [(\frac{\hat{\mu}_{11} + \hat{\mu}_{12}}{2} - \hat{\mu}_{13}) - (\frac{\hat{\mu}_{21} + \hat{\mu}_{22}}{2} - \hat{\mu}_{23})]/2$$

2.4.2   Variances of the Estimators

The variance of estimated TE comes from three subgroups, it can be expressed as:

$$var(\Delta\hat{\tau}) = var(\alpha\Delta\hat{\tau}_A + \beta\Delta\hat{\tau}_B + \gamma\Delta\hat{\tau}_U)$$

$$= \frac{\sigma^2}{N\alpha\rho(1-\rho)}\alpha^2 + \frac{\sigma^2}{N\beta\rho(1-\rho)}\beta^2 + \frac{\sigma^2}{N\alpha\rho(1-\rho)}\gamma^2$$

$$= \frac{\sigma^2}{N\rho(1-\rho)} \tag{28}$$

If the participants are equally randomized, then $var(\Delta\hat{\tau}) = \frac{4\sigma^2}{N}$.

It is not surprising that the variance of estimated TE in the fully randomised design is same as that in the conventional parallel randomised design, as equal numbers of participants are assigned to the same treatment group in these two designs. While comparing the variance in subgroups ($\frac{\sigma^2}{N\gamma\rho(1-\rho)}$, $\frac{\sigma^2}{N\beta\rho(1-\rho)}$, $\frac{\sigma^2}{N\alpha\rho(1-\rho)}$ for prefer-A, prefer-B and undecided subgroup, repectively) with that in conventional parallel randomised design, we can see that the variances in subgroups are inflated by preference factors (α, β, γ). The inflated variances arise from the reduced group size, with $N\alpha, N\beta, N\gamma$ for prefer-A, prefer-B, undecided subgroup, respectively, while the group size is N when ignoring the information of preferences.

Selection effect is obtained from two subgroups with a definite preference. Assuming equal variance for each subgroup and combining with the sample size for each subgroup, we have

$$
\begin{aligned}
var(\Delta\hat{\nu}) &= var[\frac{1}{2}((\hat{\mu}_{11} + \hat{\mu}_{21}) - (\hat{\mu}_{12} + \hat{\mu}_{22}))] \\
&= \frac{\sigma^2}{4}[\frac{1}{N\alpha\rho} + \frac{1}{N\beta\rho} + \frac{1}{N\alpha(1-\rho)} + \frac{1}{N\beta(1-\rho)}] \\
&= \frac{\sigma^2}{4N}\frac{1}{\rho(1-\rho)}(\frac{1}{\alpha} + \frac{1}{\beta})
\end{aligned}
\tag{29}
$$

and if it is designed as equally randomised, it would be $var(\Delta\hat{\nu}) = \frac{\sigma^2}{N}(\frac{1}{\alpha} + \frac{1}{\beta})$. The estimated PE involves the same items as estimated SE, thus the variance of estimated PE is the same as that of estimated SE. For the variance of the estimated second selection contrast, we could obtain it by assuming the equal variance of each subgroup

and taking the sample size of each subgroup as weight. It is shown as follows:

$$
\begin{aligned}
var(\Delta\hat{\nu}') &= var[((\hat{\mu}_{13} + \hat{\mu}_{23}) - \frac{1}{2}(\hat{\mu}_{11} + \hat{\mu}_{21} + \hat{\nu}_{12} + \hat{\nu}_{22}))/2] \\
&= \frac{\sigma^2}{4N} \frac{1}{\rho(1-\rho)}[(\frac{1}{4\alpha} + \frac{1}{4\beta}) + \frac{1}{\gamma}]
\end{aligned}
\tag{30}
$$

Again, the variance of estimated second preference contrast is same as that of estimated second selection contrast.

In the case of equal randomization, $var(\Delta\hat{\nu}') = \frac{\sigma^2}{N}(\frac{1}{4\alpha} + \frac{1}{4\beta} + \frac{1}{\gamma})$. We can see that the variance of estimated second selection contrast for the fully randomized design is same as that for the partially randomized design. Otherwise, they are not equal.

.

Chapter 3

Results

3.1   Treatment Effect

Figure 3.1 shows the relative efficiency of treatment effect for four different designs with respect to various proportions of undecided participants. The relative inverse SEs are taken as a measure of relative efficiency, where the base case is standard error of the estimated TE for the conventional trials. As shown in the Figure 3.1, relative efficiency of the estimated TE in the fully randomized design is same as that in the conventional trials. Except for the partially randomized design, relative efficiencies in other three designs are independent of the undecided rate. Relative efficiency for the two-stage design is $1/\sqrt{2}$ (Table 3.1), compared to the base case. It is worth to mention that when the proportion of undecideds is 100%, partially randomized design and fully randomized design are equivalent since all identified undecided participants are randomized to one of the two treatment. In the conventional trials, although the preferences of participants are not identified at the first stage, they are randomized in the same way as partially and fully randomized design. Regarding the two-stage

design, the choice group and random group both only include undecided participants, however, only the participants in the choice group will have been asked their preferences, whereas those in the randomized group will not. Therefore, as shown in Table 3.1, relative efficiencies of TE with respect to the conventional design, partial randomized and fully randomized design are same, with relative efficiency for the two-stage design remaining at $1/\sqrt{2}$. The decrease of relative efficiency arises from the reduced sample size in the random group. Compared to the conventional trials and fully randomized design, the two-stage design only randomizes a portion of participants. As for the partially randomized design, only undecided participants are randomized. Therefore, relative efficiency of the estimated TE for the partially randomized design depends on the undecided rate. Relative efficiency is improved with a greater portion of undecided participants. In addition, when the undecided rate is greater than 50%, relative efficiency for the partially randomized design is higher than that for the two-stage design. In summary, conventional trials and fully randomized design have higher efficiency for estimating the treatment effect than other designs. Relative efficiency of the estimated TE for the two-stage design depends on the portions of participants assigned to the random arm; and relative efficiency is relatively low for the partially randomized design when there are small portions of undecided participants.

Table 3.1: Relative efficiency for the treatment effect

proportions of undecideds	0	0.1	0.3	0.5	0.7	0.9	1
Conventional trials[*]	1	1	1	1	1	1	1
Two-stage design	0.71	0.71	0.71	0.71	0.71	0.71	0.71
Partially randomized design	NaN	0.32	0.55	0.71	0.84	0.95	1
Fully randomized design	1	1	1	1	1	1	1

[*] The base cases are SEs of the estimated treatment effects in the conventional trials
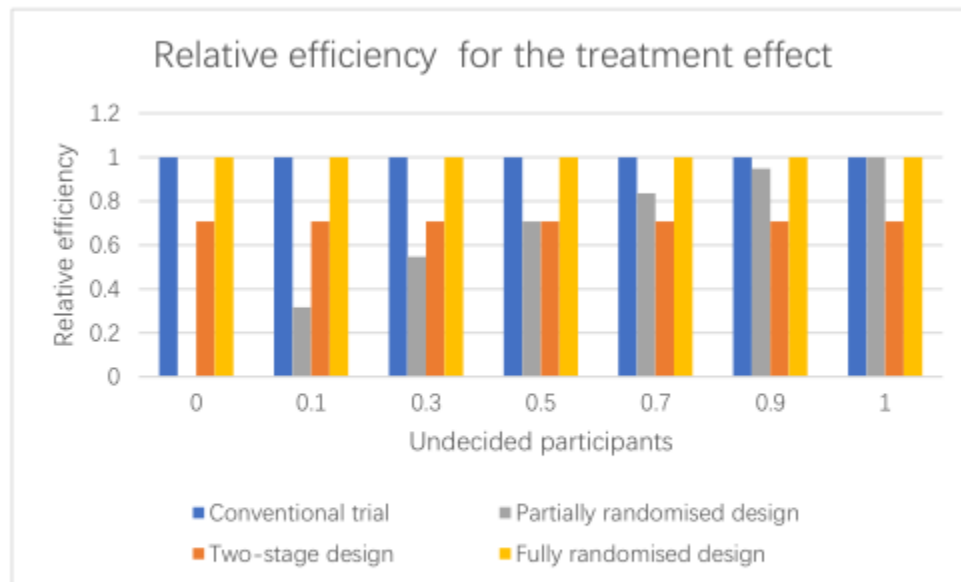
* Base cases are SEs of estimated treatment effect for conventional trials.

Figure 3.1: Relative efficiency for the treatment effect

3.2 Relative Efficiency of the First and Second Contrasts with Equal Randomization

3.2.1 Relative Efficiency of the First Contrasts

Relative efficiency under the case of equal randomization will be initially analyzed. Equal randomization means that equal numbers of participants in the random arm are assigned to the treatments. Figure 3.2 and table 3.2 show the relative efficiency of the first selection and preference contrasts for the two-stage design with a base case of 50% of undecided participants and equal preference for the treatment. It is also noted that relative preference rate for A is defined as $\frac{\alpha}{(\alpha+\beta)}$. The y-axis in the Figure 3.2 is based on logarithmic scale. As shown in Figure 3.2, as the proportion of undecideds increases, relative efficiency decreases faster. For instance, when the treatments are equally preferred, relative efficiency with no undecided participants is 4, while it is 3.24 with 10% undecideds and only 0.04 with 90% undecideds. With the proportion of undecided participants (γ) fixed, relative efficiency is greatest when the two treatments are equally preferred, and it has symmetric pattern with respect to the relative preference rate for treatment A. Moreover, if there are 100% undecided participants, then first contrasts are not estimable as the preference distribution for the treatments is not available in the choice group in this case.
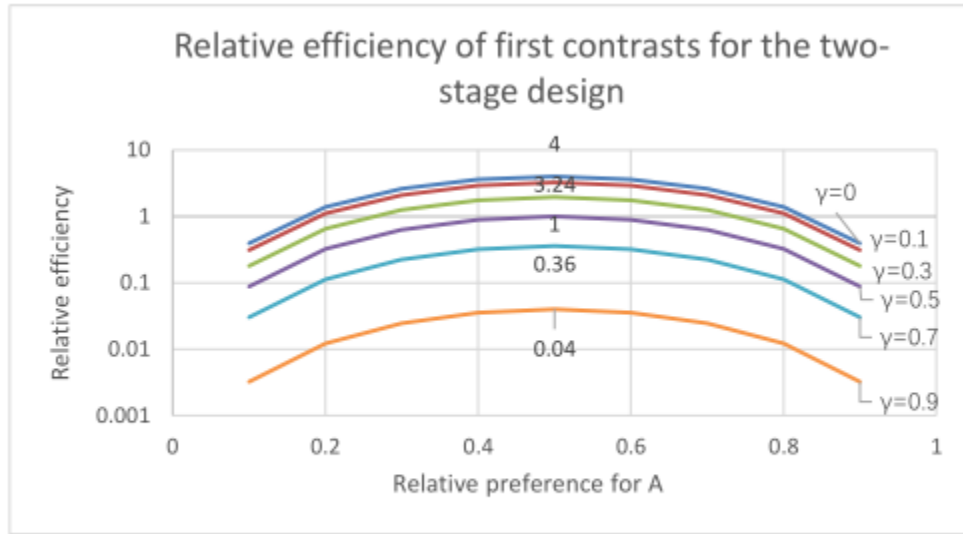
Figure 3.2: Relative efficiency of first contrasts for the two-stage design

Table 3.2: Relative efficiency of first contrasts for the two-stage design

relative preference for A	undecided rate					
	0	0.1	0.3	0.5	0.7	0.9
0.1	0.39	0.31	0.18	0.09	0.03	0.00
0.2	1.39	1.11	0.65	0.32	0.11	0.01
0.3	2.61	2.10	1.25	0.63	0.22	0.02
0.4	3.61	2.92	1.76	0.89	0.32	0.04
0.5	4.00	3.24	1.96	1.00	0.36	0.04
0.6	3.61	2.92	1.76	0.89	0.32	0.04
0.7	2.61	2.10	1.25	0.63	0.22	0.02
0.8	1.39	1.11	0.65	0.32	0.11	0.01
0.9	0.39	0.31	0.18	0.09	0.03	0.00

[*] The base case is undecided rate $\gamma = 50\%$ and equal preference for the treatments under two-stage design

To compare the designs consistently, the baseline for the two-stage design is taken as the base for both the partially randomized design and the fully randomized design.

As shown in the Figure 3.3 and Table 3.3, relative efficiencies of the selection effect with no undecided participants and 100% undecideds are not estimable for the partially randomized design since selection effect of the partially randomized design depends on the both the group of participants who have a preference and the group with no preference. Compared to the base case with a 50% undecided participants and equal preferences for the treatments under two-stage design, relative efficiency of the estimated SE for the partially randomized design is greater than 1 in all various proportions of undecided participants except in the case where the proportion of undecideds $\gamma = 90\%$ and treatments are far from equally preferred. Specifically, relative efficiency with 90% of undecideds and 10% relative preference rate for A is 0.55; and a study with equal rate of undecideds and 20% relative preference rate for A has an efficiency of 0.96. In addition, relative efficiency is not monotonically related to the rate of undecideds.Relative efficiency is maximized with 50% undecideds and drops fast when the undecided rate is far from 50%. It is also noted that relative efficiency of the selection effect for the two-stage design shows a symmetric pattern around a relative preference of 50:50.

Figure 3.3: Relative efficiency of first selection contrast for the partially randomized design

Table 3.3: Relative efficiency of first selection contrast for the partially randomized design

	undecided rate					
relative preference for A	0	0.1	0.3	0.5	0.7	0.9
0.1	NaN	1.22	2.19	2.12	1.50	0.55
0.2	NaN	1.36	2.87	3.12	2.41	0.96
0.3	NaN	1.41	3.18	3.65	2.96	1.23
0.4	NaN	1.43	3.32	3.92	3.26	1.39
0.5	NaN	1.44	3.36	4.00	3.36	1.44
0.6	NaN	1.43	3.32	3.92	3.26	1.39
0.7	NaN	1.41	3.18	3.65	2.96	1.23
0.8	NaN	1.36	2.87	3.12	2.41	0.96
0.9	NaN	1.22	2.19	2.12	1.50	0.55

[*] The base case is undecided rate $\gamma = 50\%$ and equal preference for the treatments

Figure 3.4 and Table 3.4 shows the same scenarios for the fully randomized design. Here the pattern is similar to that for the two-stage design, where relative efficiency is monotonically related to the undecided rate γ. Moreover, estimators of first contrasts could be quantified if there are no undecided participants, and they are not able to be obtained if all participants have no preference. However, relative efficiencies of the first contrasts for the fully randomized design are significantly higher than those for the two-stage design for all various undecideds rates. For instance, if the treatments are equally preferred, relative efficiency of a study with 90% of undecided participants is 0.04 with respect to the base case in the two-stage design, while it is 1.6 in the fully-randomized design.; and a study with 50% of undecided participants has a relative efficiency of 8, compared to the base case for two-stage design (relative efficiency is one).
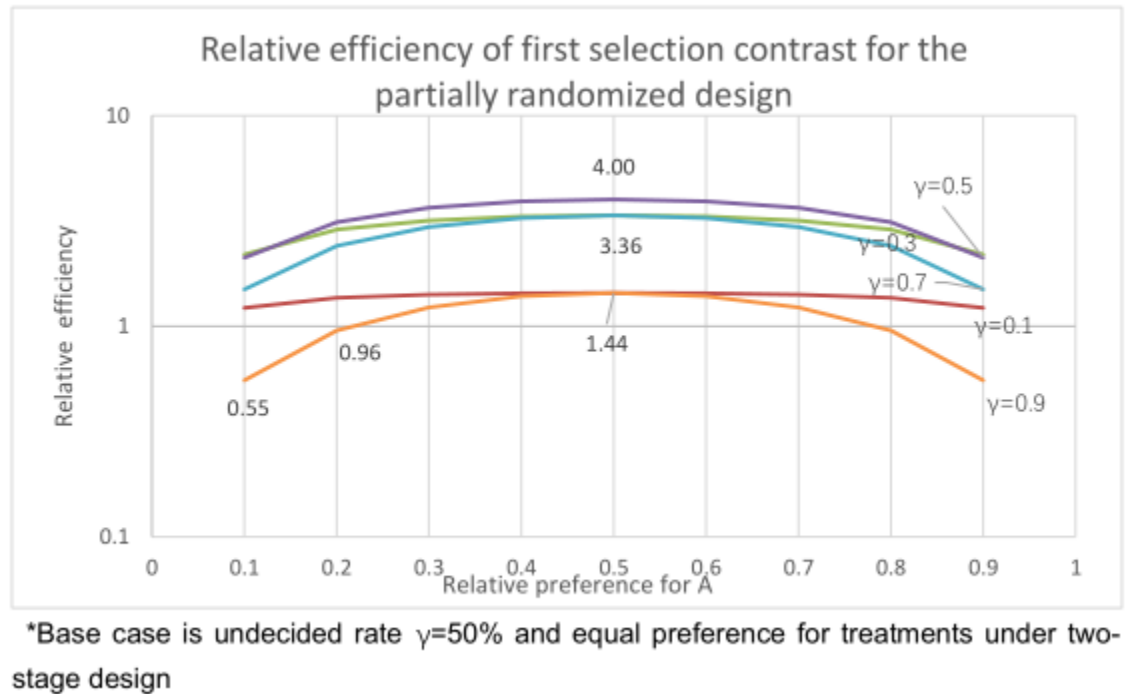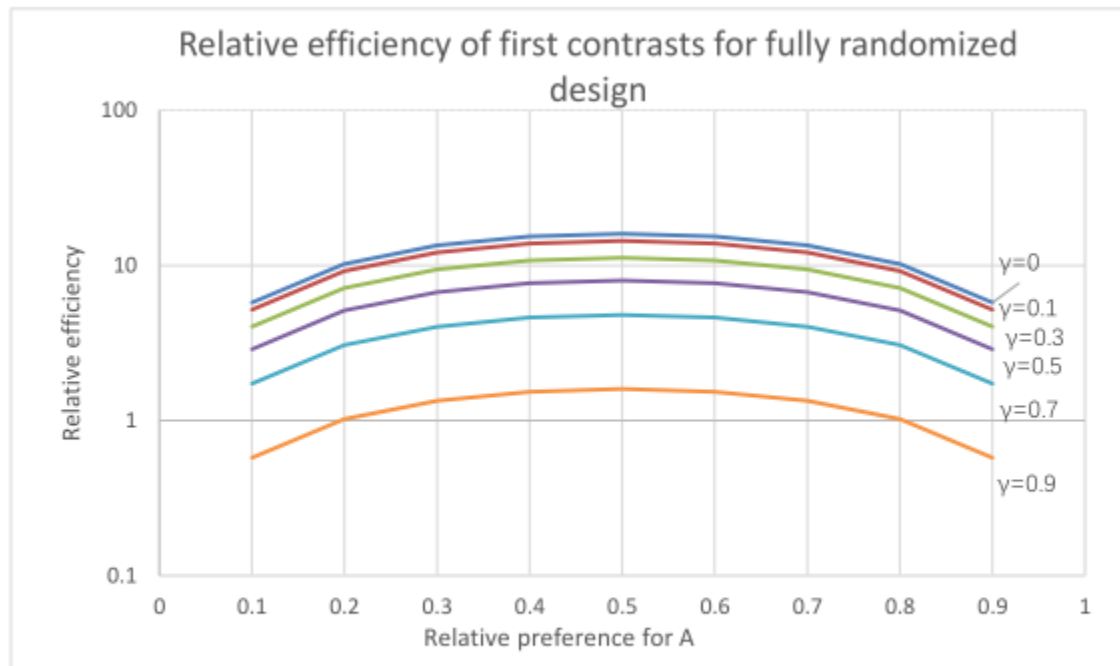
Figure 3.4: Relative efficiency of first contrasts for fully randomized design

Table 3.4: Relative efficiency of first contrasts for fully randomized design

	undecided rate					
relative preference for A	0	0.1	0.3	0.5	0.7	0.9
0.1	5.76	5.18	4.03	2.88	1.73	0.58
0.2	10.24	9.22	7.17	5.12	3.07	1.02
0.3	13.44	12.10	9.41	6.72	4.03	1.34
0.4	15.36	13.82	10.75	7.68	4.61	1.54
0.5	16.00	14.40	11.20	8.00	4.80	1.60
0.6	15.36	13.82	10.75	7.68	4.61	1.54
0.7	13.44	12.10	9.41	6.72	4.03	1.34
0.8	10.24	9.22	7.17	5.12	3.07	1.02
0.9	5.76	5.18	4.03	2.88	1.73	0.58

[*] The base case is undecided rate $\gamma = 50\%$ and equal preference for the treatments under two-stage design

To summarize, relative efficiencies of the first contrasts are generally higher in the fully randomized design and partially randomized design than those in other designs. Therefore, if selection preference (and preference effect) is (are) of greatest interest to investigators, then partially randomized design and fully randomized design might be the first choices (if ignoring other factors which might affect the experiment). Two-stage design might only be considered if the portions of undecided participants are relatively small.

3.2.2   Relative Efficiency of the Second Contrasts

For the second selection and preference contrasts, the base case is still defined by having 50% undecided participants and equally preferred treatments. Note that second contrasts measure the difference in the outcomes between participants with no preference and those with a preference. Thus, in order to estimate the second contrasts, there must exist these two types of participants (no preference vs. a specific preference) in the design. Similar to the pattern of first contrasts for the two-stage design, relative efficiencies of the second contrasts (Figure 3.5 and Table 3.5) are relatively high when there are less than half undecided patients and decrease fast when there is a larger proportion of undecideds. However, here relative efficiency with 30% of undecideds is higher than that with 10% of undecideds, when the relative preference for A ranges from 0.2 to 0.8, which means that the effect of γ is not strictly monotonic. Specifically, if the treatments are equally preferred, relative efficiency with 30% of undecideds is 1.36, while relative efficiency with 10% of undecideds is 0.88. It is also noted that when there is an extremely high preference for one treatment, the difference between relative efficiency of 10% of undecideds and

90% of undecideds is very significant. Relative efficiency is 0.005 if there are 10% of undecided participants and 90% of relative preference for the treatment A while 0.309 if there are 90% of undecided participants. As it shows, the relative efficiencies are quite low in both cases, which are not reasonable to accept. Relative efficiencies of the estimated second contrasts for the two-stage design is greater 1 only when the proportions of undecided participants are slightly less than 50% and roughly equal preference for the treatments. (Relative efficiency of 1.36 when undecided rate is 30% and the treatments are equally preferred)

Figure 3.6 shows the pattern for both the partially and fully randomized designs. As shown in chapter 2, variance of the second preference contrast is not estimable in the partially randomized design while variance of the second selection contrast for the partially randomized design is same as that for the fully randomized design when the participants in the random arm are equally randomized. In addition, the pattern here is quite different from that for the two-stage design. The relative efficiency is low when there is either an extremely small or large proportion of undecided participants. Having a 50% of undecided participants attains the highest level of efficiency. Overall, relative efficiencies of the estimated second contrasts in the partially randomized design and fully randomized design are significantly greater than that in the two-stage design. As shown in Figure 3.6 and Table 3.6, relative efficiency is significantly greater than 1 when relative preference for A ranges from 0.2 to 0.8 in the partially and fully randomized designs, while relative efficiency is slightly above 1 only under a rate of undecideds 30% and a relative preference rate not far away from 50%.
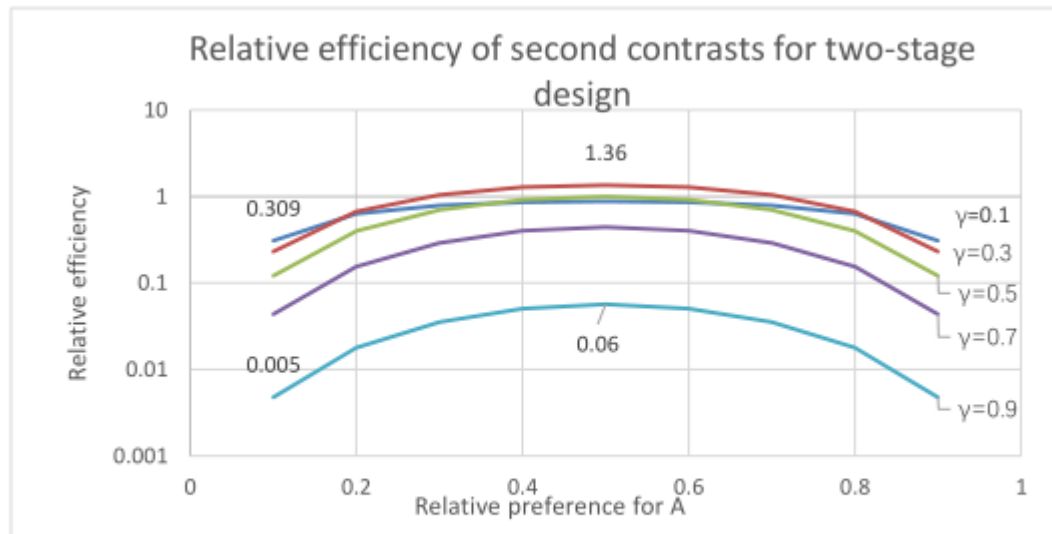
Figure 3.5: Relative efficiency of second contrasts for two-stage design

Table 3.5: Relative efficiency of second contrasts for two-stage design

	undecided rate				
relative preference for A	0.1	0.3	0.5	0.7	0.9
0.1	0.31	0.23	0.12	0.04	0.00
0.2	0.63	0.67	0.40	0.15	0.02
0.3	0.79	1.05	0.70	0.29	0.04
0.4	0.86	1.28	0.92	0.40	0.05
0.5	0.88	1.36	1.00	0.44	0.06
0.6	0.86	1.28	0.92	0.40	0.05
0.7	0.79	1.05	0.70	0.29	0.04
0.8	0.63	0.67	0.40	0.15	0.02
0.9	0.31	0.23	0.12	0.04	0.00

[*] The base case is undecided rate $\gamma = 50\%$ and equal preference for the treatments under two-stage design

38
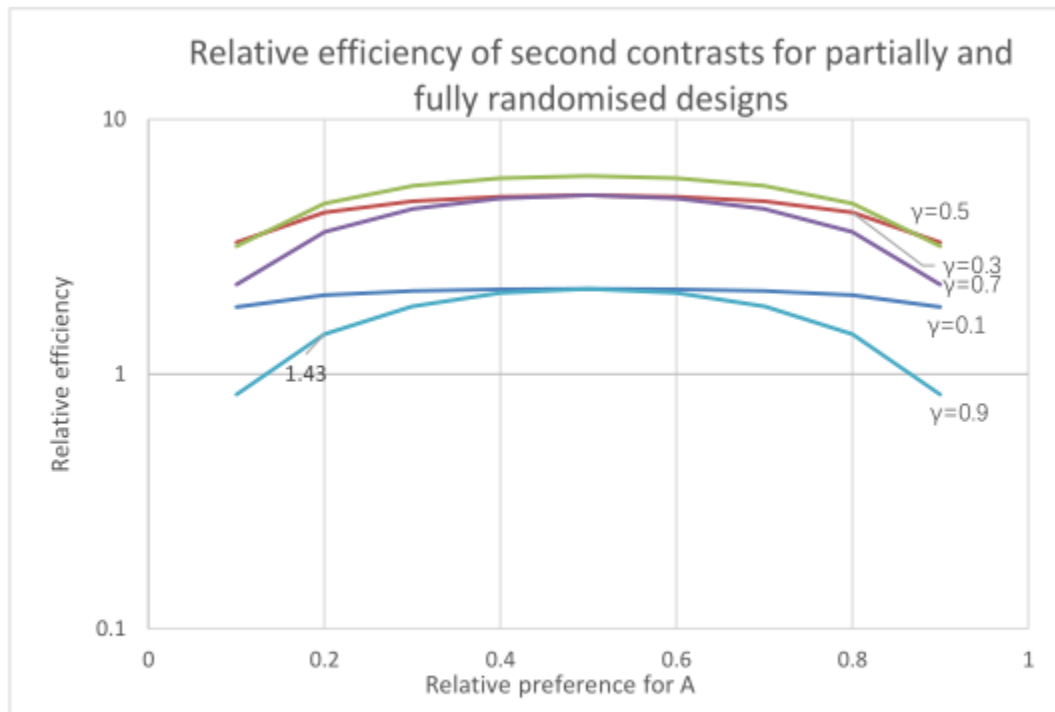
Figure 3.6: Relative efficiency of second contrasts for partially and fully randomised designs

In summary, if investigators are interested in the relationship between the undecided participants and those who have a preference, (i.e., second contrasts), then partially randomized or fully randomized designs might be taken as priorities.

Table 3.6: Relative efficiency of second contrasts for partially and fully randomised designs

	undecided rate				
relative preference for A	0.1	0.3	0.5	0.7	0.9
0.1	1.83	3.29	3.18	2.25	0.83
0.2	2.04	4.31	4.68	3.62	1.43
0.3	2.12	4.77	5.48	4.45	1.84
0.4	2.15	4.98	5.88	4.90	2.08
0.5	2.16	5.04	6.00	5.04	2.16
0.6	2.15	4.98	5.88	4.90	2.08
0.7	2.12	4.77	5.48	4.45	1.84
0.8	2.04	4.31	4.68	3.62	1.43
0.9	1.83	3.29	3.18	2.25	0.83

[*] The base case is undecided rate $\gamma = 50\%$ and equal preference for the treatments

3.3 Relative Efficiency of the First and Second Contrasts with Unequal Randomization

3.3.1 Relative Efficiency of the First Contrasts

The values shown in Figure 3.7 and Table 3.7 are based on 1:2 randomization, which means 1/3 of participants in the random arm are randomized to treatment A and the rest to treatment B. As shown in the Figure 3.7, the overall pattern here is quite similar to the pattern with equal randomization. One noticeable difference is that the curves are not symmetric when participants are not equally randomized. For instance, when the undecided rate is 30%, relative efficiency with 40% of relative preference for A is 1.95 while relative efficiency with 60% of relative preference for A is 1.64. With 1 vs. 2 randomization, relative efficiency is improved when there are a relatively greater portion of participants preferring treatment B. In addition, with unequal randomization, relative efficiency is not maximized at the point with equal

preference for the treatments for a fixed undecided rate. For instance, when there are 50% of undecided participants, the relative preference rate for A where relative efficiency is maximized is 45.2% rather than 50%. As we can see, the difference is relatively small, with 4.8% smaller than the 50%.
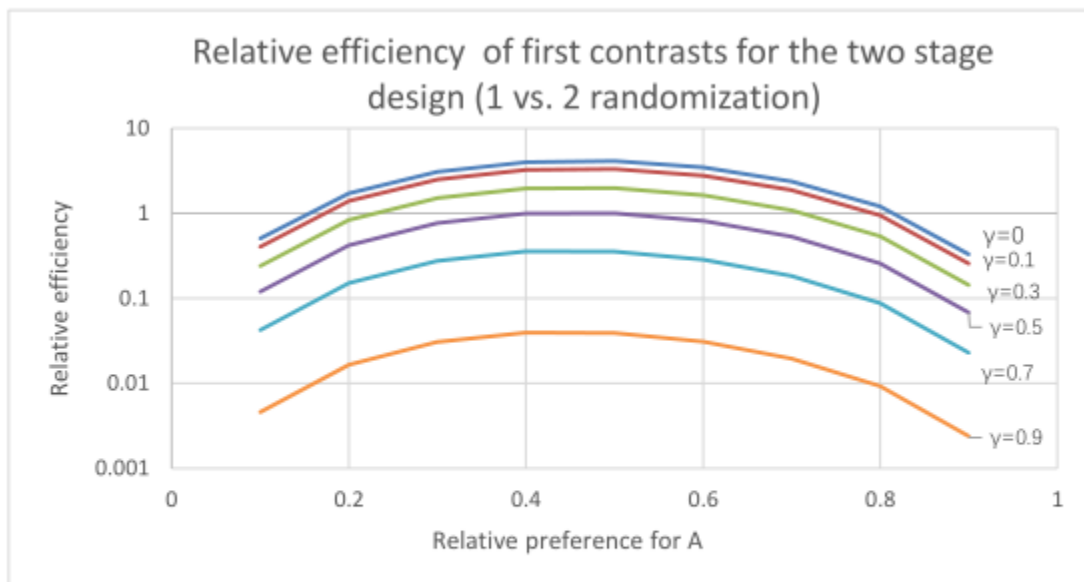
*Base case is undecided rate γ=50% and equal preference for the treatments under two-stage design.1vs.2 randomization means that 1/3 of participants in the random arm are randomized to treatment A and the rest to treatment B.

Figure 3.7: Relative efficiency of first contrasts for the two stage design (1 vs. 2 randomization)

For the partially randomized design, (Figure 3.8 and Table 3.8) relative efficiency still shows a symmetric pattern, which is similar to that for the equal randomization. A very subtle difference is that when the treatments are equally preferred, relative efficiency is maximized if there are 50% of undecided participants with equal random-ization (where relative efficiency is 2.029), but it is 51.47% of undecided participants

Table 3.7: Relative efficiency of first contrasts for the two stage design (1 vs. 2 randomization)

relative preference for A	undecided rate					
	0	0.1	0.3	0.5	0.7	0.9
0.1	0.51	0.40	0.24	0.12	0.04	0.00
0.2	1.72	1.39	0.83	0.42	0.15	0.02
0.3	3.08	2.49	1.51	0.77	0.28	0.03
0.4	3.99	3.23	1.95	0.99	0.36	0.04
0.5	4.12	3.32	1.98	1.00	0.36	0.04
0.6	3.48	2.78	1.64	0.81	0.29	0.03
0.7	2.37	1.88	1.09	0.53	0.18	0.02
0.8	1.20	0.94	0.54	0.26	0.09	0.01
0.9	0.33	0.26	0.14	0.07	0.02	0.00

[*] The base case is undecided rate $\gamma = 50\%$ and equal preference for the treatments under two-stage design. 1 vs.2 randomization means that 1/3 of participants in the random arm are randomized to treatment A and the rest to treatment B.

with unequal randomization (where relative efficiency is 2.030).
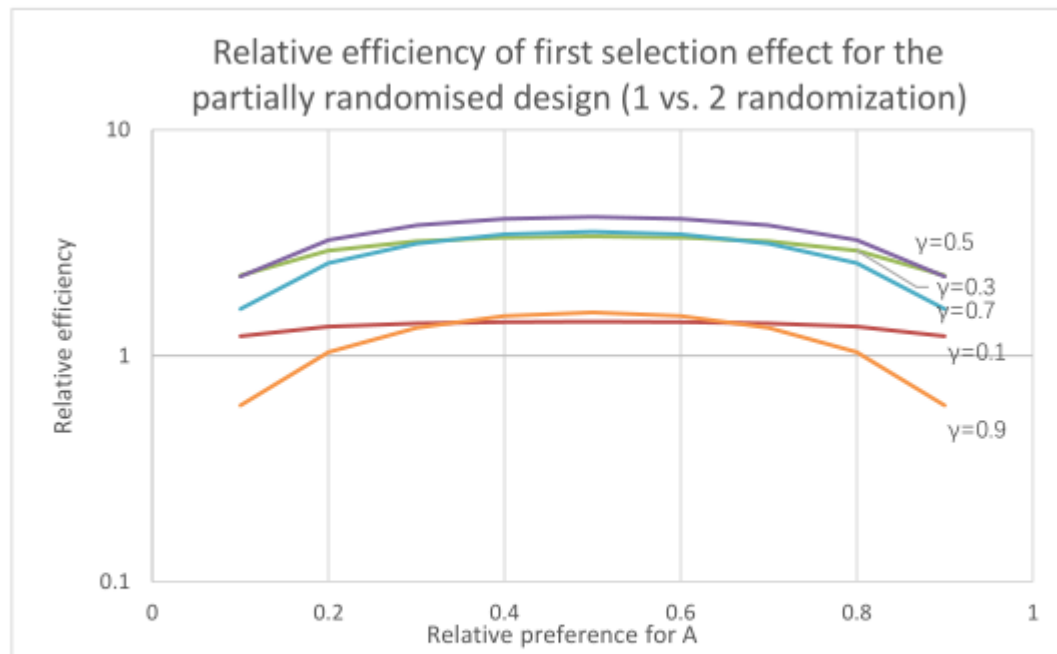
Relative efficiency of first selection effect for the partially randomised design (1 vs. 2 randomization)

* Base case is undecided rate γ=50% and equal preference for the treatments under two-stage design. 1vs.2 randomization means that 1/3 of participants in the random arm are randomized to treatment A and the rest to treatment B.

Figure 3.8: Relative efficiency of first selection effect for the partially randomised design (1 vs. 2 randomization)

As for the fully randomized design, the overall pattern (Figure 3.9 and Table 3.9) is same as that of equal randomization, where the curves are still symmetric. The relative efficiency is lost as the proportion of undecided participants γ increases. As shown in the equation 27 for Chapter 2, the variances of estimated SE and PE are proportional to the randomization ratio ρ, which can be inferred that ρ has no impact on the shape of curve. As found for the partially randomized design, the pattern with respect to the relative efficiency for the fully randomized design is not influenced by unequal allocation.

Table 3.8: Relative efficiency of first selection effect for the partially randomised design (1 vs. 2 randomization)

relative preference for A	undecided rate					
	0	0.1	0.3	0.5	0.7	0.9
0.1	NaN	1.22	2.27	2.24	1.61	0.60
0.2	NaN	1.35	2.93	3.26	2.57	1.04
0.3	NaN	1.39	3.21	3.78	3.14	1.33
0.4	NaN	1.41	3.34	4.04	3.45	1.50
0.5	NaN	1.42	3.38	4.12	3.54	1.56
0.6	NaN	1.41	3.34	4.04	3.45	1.50
0.7	NaN	1.39	3.21	3.78	3.14	1.33
0.8	NaN	1.35	2.93	3.26	2.57	1.04
0.9	NaN	1.22	2.27	2.24	1.61	0.60

[*] The base case is undecided rate $\gamma = 50\%$ and equal preference for the treatments under two-stage design. 1 vs.2 randomization means that 1/3 of participants in the random arm are randomized to treatment A and the rest to treatment B.

*Base case is undecided rate γ=50% and equal preference for the treatments under two-stage design. 1vs.2 randomization means that 1/3 of participants in the random arm are randomized to treatment A and the rest to treatment B.
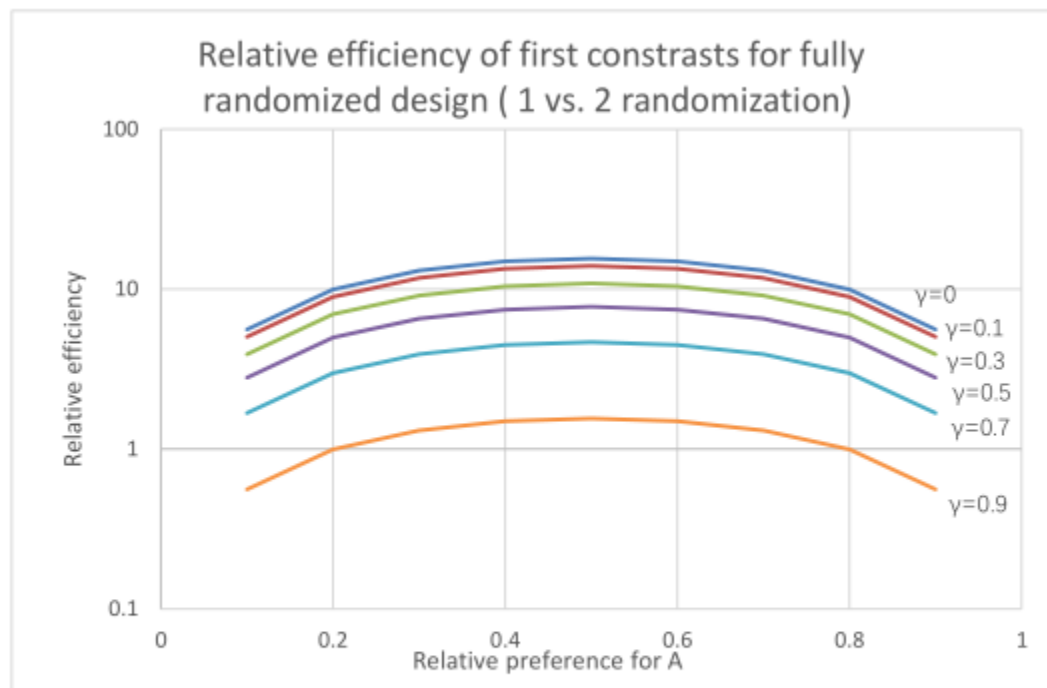
Figure 3.9: Relative efficiency of first constrasts for fully randomized design (1 vs. 2 randomization)

Table 3.9: Relative efficiency of first constrasts for fully randomized design (1 vs. 2 randomization)

relative preference for A	undecided rate					
	0	0.1	0.3	0.5	0.7	0.9
0.1	5.60	5.04	3.92	2.80	1.68	0.56
0.2	9.96	8.96	6.97	4.98	2.99	1.00
0.3	13.07	11.76	9.15	6.53	3.92	1.31
0.4	14.93	13.44	10.45	7.47	4.48	1.49
0.5	15.56	14.00	10.89	7.78	4.67	1.56
0.6	14.93	13.44	10.45	7.47	4.48	1.49
0.7	13.07	11.76	9.15	6.53	3.92	1.31
0.8	9.96	8.96	6.97	4.98	2.99	1.00
0.9	5.60	5.04	3.92	2.80	1.68	0.56

[*] The base case is undecided rate $\gamma = 50\%$ and equal preference for the treatments under two-stage design. 1 vs.2 randomization means that 1/3 of participants in the random arm are randomized to treatment A and the rest to treatment B.

3.3.2　Relative Efficiency of the Second Contrasts

Relative efficiency of second contrasts for two-stage design (1 vs. 2 randomization)

*Base case is undecided rate γ=50% and equal preference for the treatments under two-stage design. 1vs.2 randomization means that 1/3 of participants in the random arm are randomized to treatment A and the rest to treatment B.
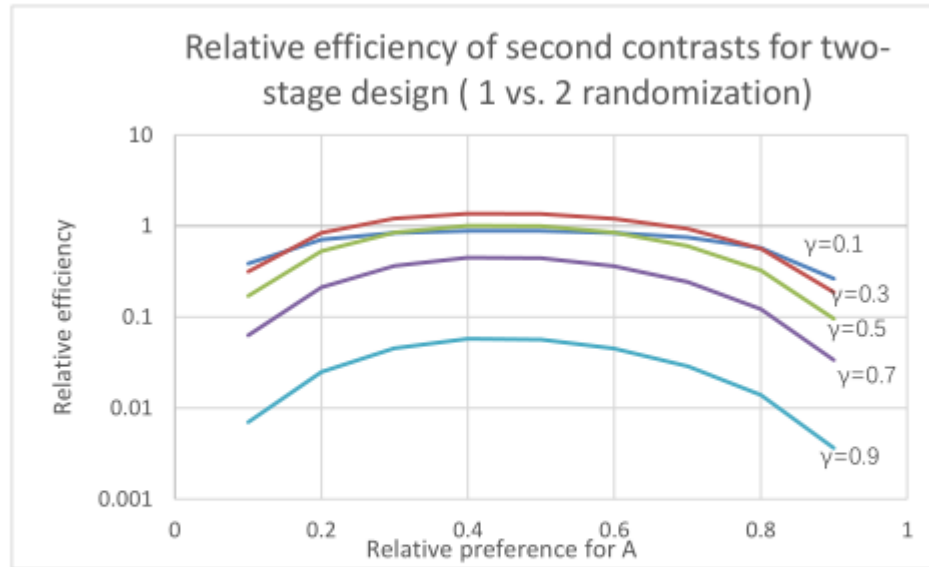
Figure 3.10: Relative efficiency of second contrasts for two-stage design (1 vs. 2 randomization)

Figure 3.10 presents an unsymmetrical pattern for the relative efficiency of second contrasts for the two-stage design, compared to Figure 3.5 with equal allocation. When the randomization rate is 1/3 (1/3 of participants assigned to treatment A, and the rest to treatment B), relative efficiency is lost if more participants are allocated to treatment B when there are greater portion of participants preferring treatment A. For example, when there are 30% undecideds, relative efficiency with a 20% relative preference rate for A is 0.84, while it is 0.57 with a 80% relative preference rate for A. (Table 3.10) Moreover, instead of being maximized at equal relative preference for the

Table 3.10: Relative efficiency of second contrasts for two-stage design (1 vs. 2 randomization)

relative preference for A	undecided rate				
	0.1	0.3	0.5	0.7	0.9
0.1	0.39	0.32	0.17	0.06	0.01
0.2	0.71	0.84	0.53	0.21	0.03
0.3	0.84	1.21	0.85	0.36	0.05
0.4	0.89	1.37	1.01	0.45	0.06
0.5	0.88	1.36	1.00	0.44	0.06
0.6	0.84	1.21	0.85	0.36	0.05
0.7	0.75	0.93	0.60	0.24	0.03
0.8	0.57	0.57	0.33	0.12	0.01
0.9	0.26	0.19	0.10	0.03	0.00

[*] The base case is undecided rate $\gamma = 50\%$ and equal preference for the treatments under two-stage design. 1 vs.2 randomization means that 1/3 of participants in the random arm are randomized to treatment A and the rest to treatment B.

treatments, relative efficiency reaches the highest level when the relative preference for the treatment A is 44.2%, which is still a relatively small difference. In addition, the undecided rate has a significant effect on efficiency especially when it is greater than 50%. For instance, if the treatments are equally preferred, relative efficiency with 30% of undecided participants is 1.36 and it is only 0.44 when there is 70% of undecided participants. Similar to the scenario with equal randomization, relative efficiency of the second contrasts for the two-stage design is only greater than one when there is slightly less than 50% of undecided participants, with 1.36 for an equal relative preference for the treatments.
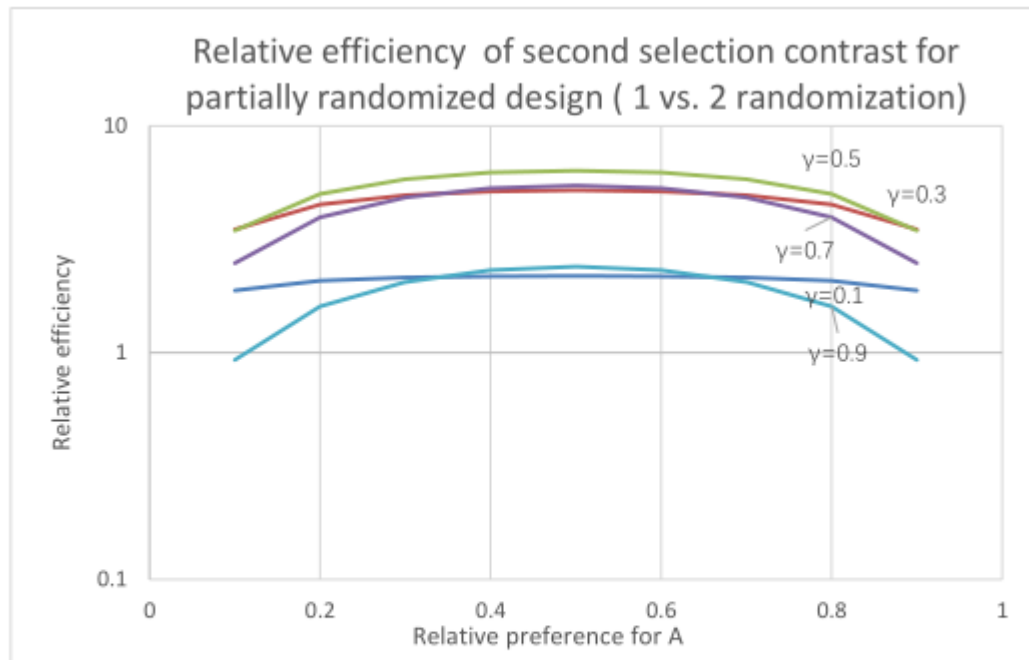
*Base case is undecided rate γ=50% and equal preference for the treatments under two-stage design. 1vs.2 randomization means that 1/3 of participants in the random arm are randomized to treatment A and the rest to treatment B.

Figure 3.11: Relative efficiency of second selection contrast for partially randomized design (1 vs. 2 randomization)

Table 3.11: Relative efficiency of second selection contrast for partially randomized design (1 vs. 2 randomization)

| | undecided rate | | | | |
relative preference for A	0.1	0.3	0.5	0.7	0.9
0.1	1.88	3.50	3.46	2.48	0.93
0.2	2.08	4.51	5.02	3.96	1.60
0.3	2.15	4.95	5.83	4.84	2.05
0.4	2.18	5.15	6.23	5.32	2.31
0.5	2.18	5.21	6.35	5.47	2.40
0.6	2.18	5.15	6.23	5.32	2.31
0.7	2.15	4.95	5.83	4.84	2.05
0.8	2.08	4.51	5.02	3.96	1.60
0.9	1.88	3.50	3.46	2.48	0.93

[*] The base case is undecided rate $\gamma = 50\%$ and equal preference for the treatments under two-stage design. 1 vs.2 randomization means that $1/3$ of participants in the random arm are randomized to treatment A and the rest to treatment B.
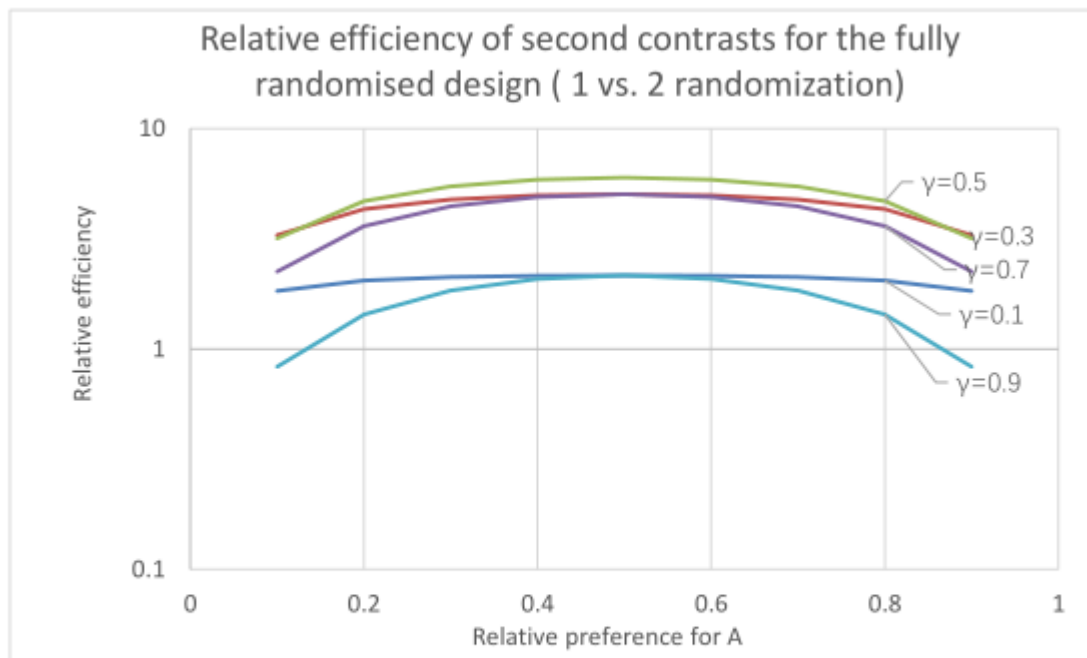

*Base case is undecided rate γ=50% and equal preference for the treatments under two-stage design. 1vs.2 randomization means that 1/3 of participants in the random arm are randomized to treatment A and the rest to treatment B.

Figure 3.12: Relative efficiency of second contrasts for fully randomized design (1 vs. 2 randomization)

Table 3.12: Relative efficiency of second contrasts for fully randomized design (partially randomized design) (1 vs. 2)

relative preference for A	undecided rate				
	0.1	0.3	0.5	0.7	0.9
0.1	1.83(1.88)	3.29(3.50)	3.18(3.46)	2.25(2.48)	0.83(0.93)
0.2	2.04(2.15)	4.31(4.51)	4.68(5.02)	3.62(3.96)	1.43(1.60)
0.3	2.12(2.18)	4.77(4.95)	5.48(5.83)	4.45(4.84)	1.84(2.05)
0.4	2.15(2.18)	4.98(5.15)	5.88(6.23)	4.90(5.32)	2.08(2.31)
0.5	2.16(2.18)	5.04(5.21)	6.00(6.35)	5.04(5.47)	2.16(2.40)
0.6	2.15(2.18)	4.98(5.15)	5.88(6.23)	4.90(5.32)	2.08(2.31)
0.7	2.12(2.15)	4.77(4.95)	5.48(5.83)	4.45(4.84)	1.84(2.05)
0.8	2.04(2.08)	4.31(4.51)	4.68(5.02)	3.62(3.96)	1.43(1.60)
0.9	1.83(1.88)	3.29(3.50)	3.18(3.46)	2.25(2.48)	0.83(0.93)

[*] The values shown in bracket are the relative efficiency for partially randomized design. The base case is undecided rate $\gamma = 50\%$ and equal preference for the treatments under two-stage design. 1 vs.2 randomization means that 1/3 of participants in the random arm are randomized to treatment A and the rest to treatment B.

Unequal allocation has a trivial impact on the shape of curve as shown in Figure 3.11 and Figure 3.12. Specifically, Figure 3.6 shows equal relative efficiency of second contrasts for the partially randomized design and fully randomized design if participants are equally randomized. In the case of unequal randomization (Table 3.11 and Table 3.12), there is a subtle difference between the two values. For instance, if the treatments are equally preferred and the proportion of undecided participants is fixed at 50%, then the relative efficiency is 6 for the fully randomized design, whereas it is 6.35 for the partially randomized design.

3.4 Comparison within Each Design Under Equal and Unequal Randomization

3.4.1 Comparison within Each Design for the First Selection and Preference Contrasts

Taking into consideration of the practical issue, we choose the proportion of undecided participants $\gamma = 30\%$ as an example for illustration purpose. When comparing within the design for the equal and unequal randomization, the equally randomized case is taken as the base.

Figure 3.13 presents a mirror symmetric pattern between the 1:2 randomization and 2:1 randomization for the two-stage design. Efficiency could be improved with unequal allocation when either treatment is preferred by most participants who a have preference. If one treatment is significantly more desirable by participants, then a gain in efficiency could be achieved by allocating more participants to that treatment. In contrast, efficiency is lost when more participants assigned to the treatment which is less desirable. For instance, if the relative preference rate for the treatment A is 10%, relative efficiency is 1.11 with 1:2 randomization and it is 0.85 with 2:1 randomization.

Figure 3.13: Relative efficiency of first contrasts for the two-stage design with 30% of undecideds (equal randomization is taken as base)

The pattern for the partially randomized design is quite different from that for the two-stage design. As shown in the Figure 3.14, we have the same precision with 1:2 randomization and 2:1 randomization. In addition, the precision of unequal randomization is always below that of equal randomization. The relative efficiency for unequal randomization changes as relative preference rate varies, which means the precision for the 1 vs. 2 (or 2 vs. 1) is not proportional to that for the equally randomized case. And compared with the precision for the equal randomization, the precision in the case of unequal randomization reaches the lowest level when treatments are equally preferred.

Figure 3.14: Relative efficiency of first selection contrast for the partially randomised design with 30% of undecideds (equal randomization is taken as base)

As for the fully randomized design, the precisions of 1:2 and 2:1 randomization are equal, and they are proportional to that of 1:1 randomization (Figure 3.15). Overall, unequal randomization leads to a reduced efficiency of first contrasts. Referred to formulas (29) in Chapter 2, the ratio is exactly $\sqrt{4\rho(1-\rho)}$.
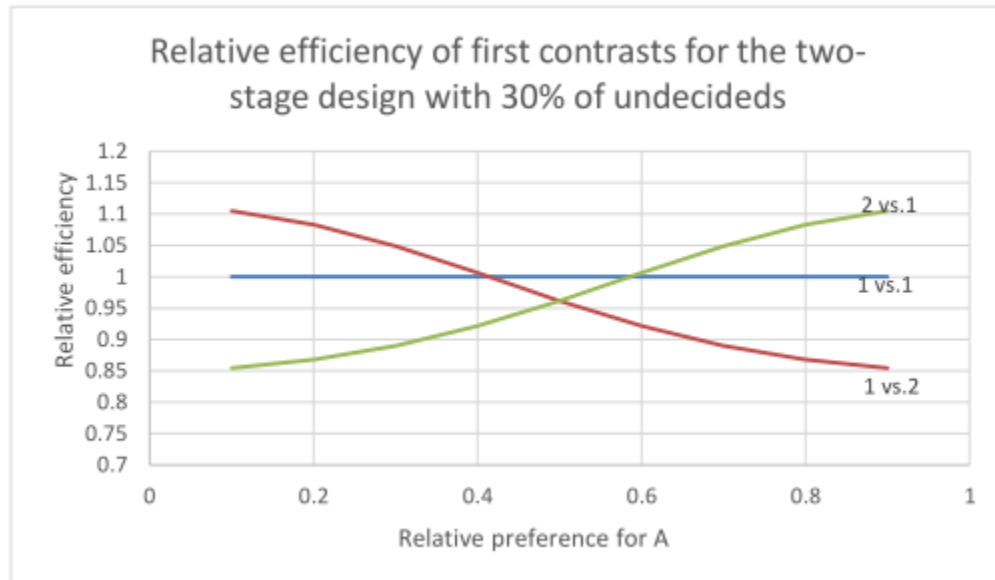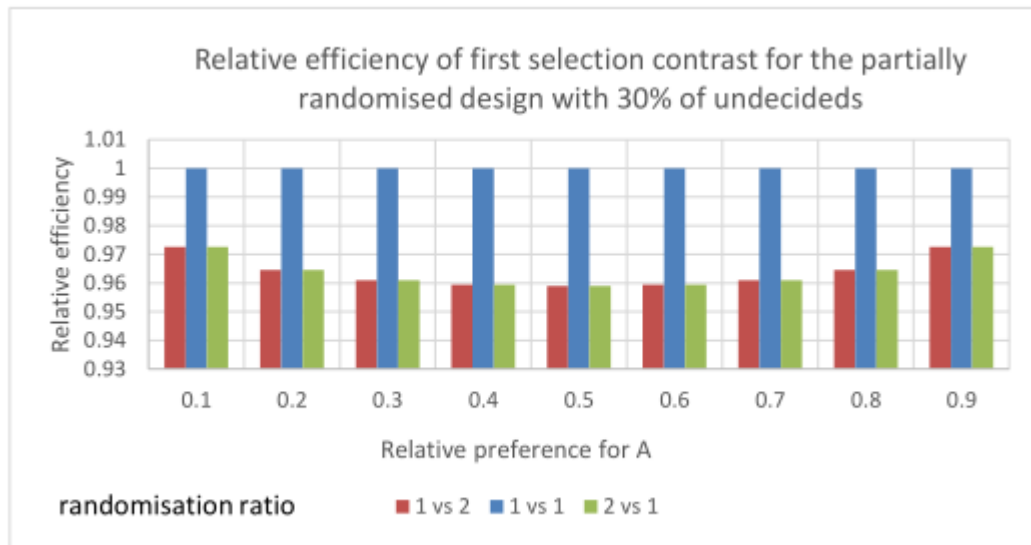
Figure 3.15: Relative efficiency of first contrasts for the fully randomised design with 30% of undecideds (equal randomization is taken as base)

3.4.2 Comparison within Each Design for the Second Selection and Preference Contrasts

As for the second contrasts, we still use the equal randomization as the base case. Figure 3.16 presents a similar pattern as found in the first contrasts. The precision of the estimated second selection and preference contrasts under the unequal randomization is lost as the relative preference rate is around 50%. Specifically, relative efficiency is 0.95 with equal preference to the treatments, compared to base case; However, unequal randomization brings benefit when one treatment is much more popular among participants who have a preference. For instance, relative efficiency is 10% (1.1) greater than the base case with 10% relative preference rate.

Figure 3.16: Relative efficiency of second contrasts for the two-stage design with 30% of undecideds (equal randomization is taken as base)

For the partially randomized design and fully randomized design, patterns for the second contrasts are also similar as we find in the first contrasts. In the partially randomized design, the precision is always lower than that of equal randomization but increases as the relative preference ratio is further away from 0.5 (Figure 3.17).
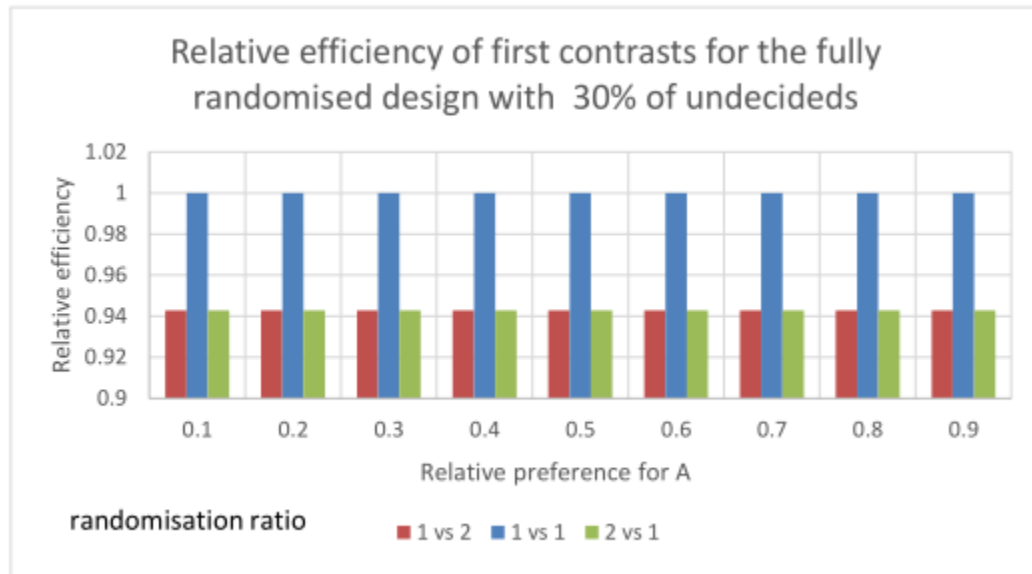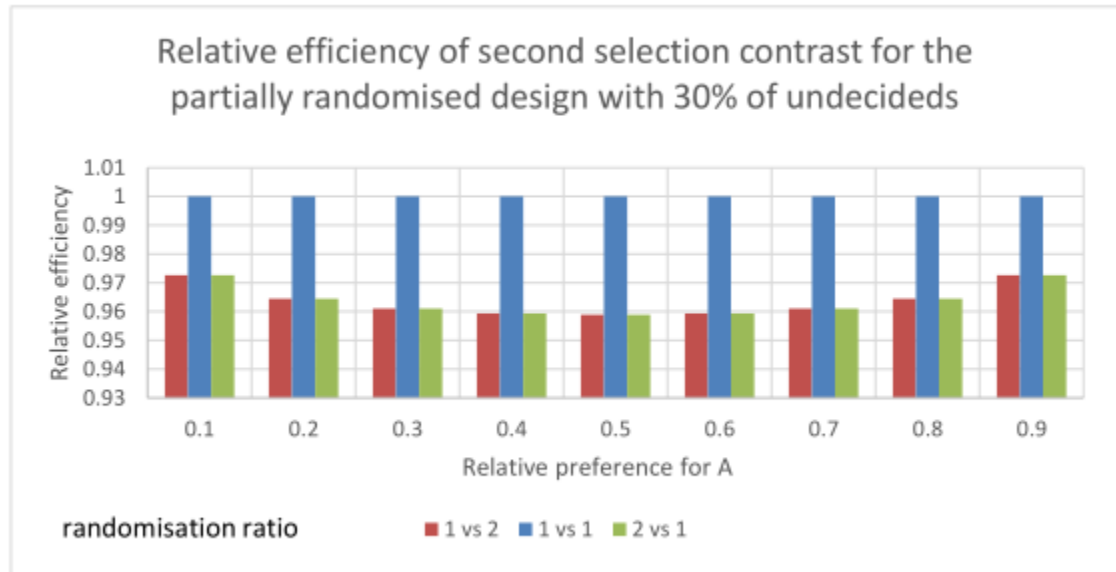
Figure 3.17: Relative efficiency of second selection contrast for the partially randomised design with 30% of undecideds (equal randomization is taken as base)

As regards the fully randomized design, the relative efficiencies of second contrasts are proportional to those of the base case, with the ratio of $\sqrt{4\rho(1-\rho)}$ (Figure 3.18).
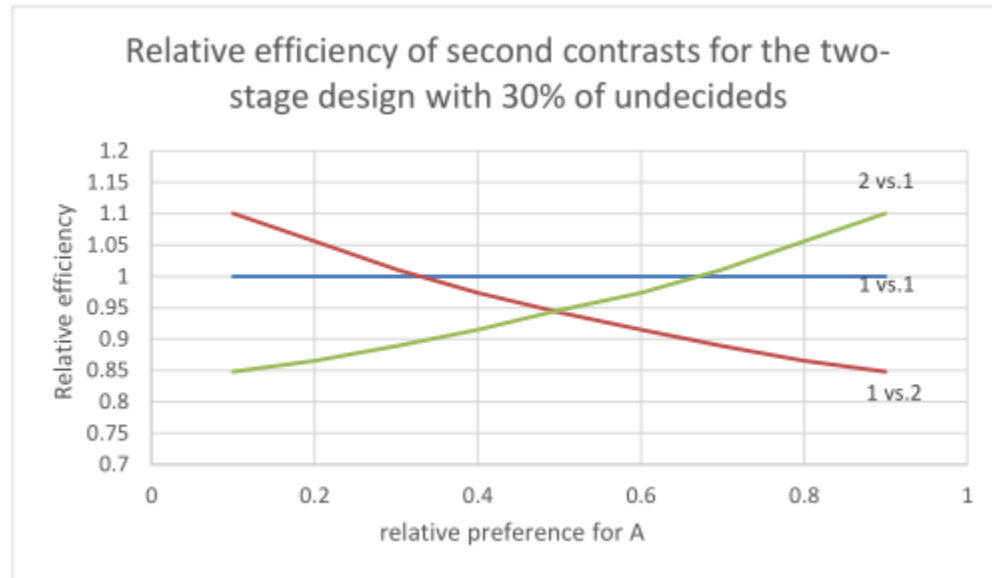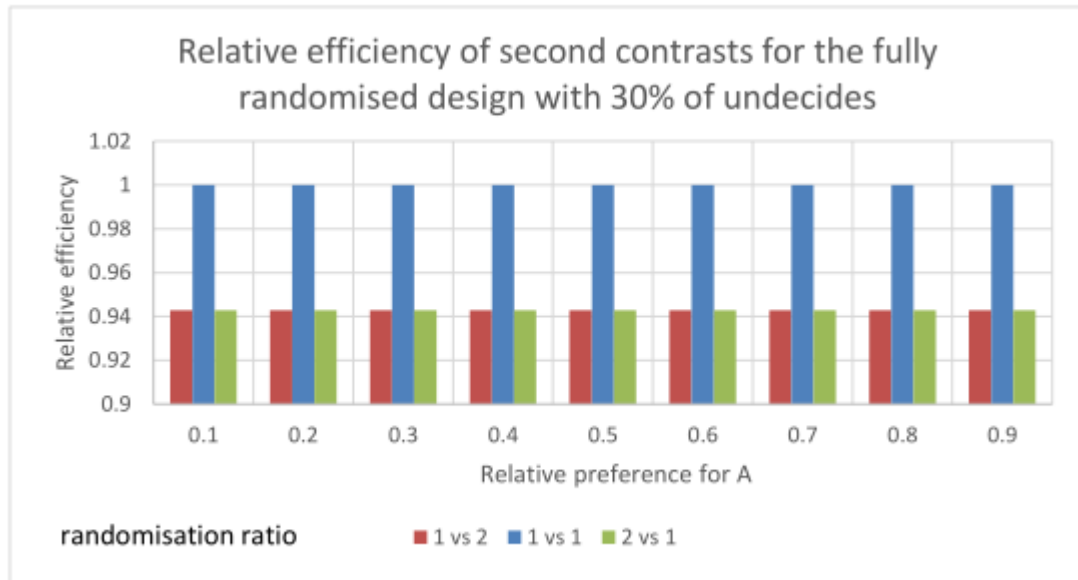
Figure 3.18: Relative efficiency of second contrasts for the fully randomised design with 30% of undecideds (equal randomization is taken as base)

Chapter 4

Practical Example

In this Chapter, we will use data from Cooper's study which evaluated different treatments for heavy menstrual bleeding, to analyze the performances of the different designs under various scenarios.

In Cooper's study, there were two treatments available. One was the medical treatment, and the alternative was surgical treatment. The total sample size was 227. 130 out of 227 participants were allocated to the choice group, and the rest of them (97) were allocated to the random group and then equally randomized to one of the two treatments. In the choice group (130), there was a total of 90 participants with no preference and they were equally randomized to the treatment. Among those who have a definite preference, 19 out of 40 chose medical treatment and 21 preferred surgical treatment. To summarize, the undecided rate was 90/130 (69.2%), the proportion of participants who have a preference was (19+21)/130 (30.8%), where the preference rates for the medical and surgical treatments are 19/130 (14.6%) and 21/130 (16.2%), respectively. (Table 4.1)

Table 4.1: Summary of Cooper's study

treatment received		Choice arm			Random arm
		chose medical	chose surgical	undecided	
Medical	Sample size	19		45	49
	Mean	16.6		18.4	17.2
	SD[*]	8.7		10.7	5.2
Surgical	Sample size		21	45	48
	Mean		5.9	4.3	5.1
	SD[*]		7.2	5.2	7.7

[*] Standard deviation

For simplicity, we assume that every subgroup has the constant variance. The pooled standard deviation ($\sigma = 7.59$) over the all data was taken as the constant standard deviation for each subgroup.

Inserting the values into formulas which are introduced in Chapter 2, we can get estimated effects and corresponding variances for various designs. As shown in Table 4.2, estimated TE was same in the conventional design, two-stage design and fully randomized design, which was due to the unbiasedness of the estimated TE in these three designs. It also showed a larger TE effect in the partially randomized design, which was 16.5% greater than the value in other designs. But this estimator of TE was biased since it was estimated from the undecided subgroup only. Regarding the two-stage design, first selection contrast was about 1/4 of treatment effect, showing that among those who were given the choice, the participants who preferred surgery had better results than those who preferred medical treatment (the observed outcomes are measured by the bleeding score, so the lower value is better than the higher one). The estimated values were smaller for the preference effect and second selection contrast,

but they have the same direction as the TE. It means that there was a larger treatment effect among those participants who would select medical treatment if possible. In addition, the participants who had a specific preference show better outcomes than those who have no preference. Second preference contrast had the opposite direction to other effects, which means that there was a larger TE among those participants with no preference. In the partially randomized design, first contrast PE was not estimable and neither was second contrast PE. A noticeable difference relative to the other designs was that the estimated SE in the partially randomized design had the opposite direction to other designs. The negative sign indicates a larger selection effect within participants who receive surgical treatment.

Table 4.3 shows the standard errors (SEs) of estimable effects for various designs. Standard errors of estimated TE in the conventional design and fully randomized design are equal and the smallest among the designs. But these two designs have practical issues. Only treatment effect can be estimated in the conventional design. For the fully randomized design, participants with a definite preference may be randomized to the treatment which they did not prefer.

Precisions of estimated SE in the partially randomized design and fully randomized design were much better than the two-stage design. These two designs also had the same precision of estimated second selection contrast, which was less than 1/3 of the standard error of two-stage design, which was due to the relatively high proportion of undecideds (69.2%). Overall, the partially randomized design had good precisions of estimable effects in the Cooper's example.

Table 4.2: Estimated effects for the Cooper's example

	$\Delta\tau$	$\Delta\nu$	$\Delta\pi$	$\Delta\nu'$	$\Delta\pi'$
Conventional	12.1	NaN	NaN	NaN	NaN
Two-stage	12.1	3.053	0.947	0.574	-3.226
Partially randomised	14.1	-3.4	NaN	-0.1	NaN
Fully randomized	12.1	3.053	0.947	0.574	-3.226

Table 4.3: Standard errors of the estimated effects for the Cooper's example

	$\Delta\tau$	$\Delta\nu$	$\Delta\pi$	$\Delta\nu'$	$\Delta\pi'$
Conventional	1.01	NaN	NaN	NaN	NaN
Two-stage	1.54	6.64	6.64	3.62	3.62
Partially randomized	1.21	2.18	NaN	1.09	NaN
Fully randomized	1.01	1.82	1.82	1.09	1.09

Table 4.4: Relative efficiencies of the estimated effects for the Cooper's example

	$\Delta\nu$	$\Delta\pi$	$\Delta\nu'$	$\Delta\pi'$
Two-stage	1.00	1.00	1.00	1.00
Partially randomized	3.04	NaN	3.32	NaN
Fully randomized	3.65	3.65	3.32	3.32

[*] The base cases are the estimated first contrasts and second contrasts of the two-stage design

Table 4.5: Relative efficiencies of the estimated effects with undecided rate 10%

	$\Delta\nu$	$\Delta\pi$	$\Delta\nu'$	$\Delta\pi'$
Two-stage	2.93	2.93	1.45	1.45
Partially randomized	1.98	NaN	2.16	NaN
Fully randomized	6.25	6.25	2.16	2.16

[*] the relative preference rate for two treatments remain unchanged (47.4% vs. 52.6%), and the base cases are the estimated first contrasts and second contrasts of the two-stage design in the Cooper's example

Table 4.6: Relative efficiencies of the estimated effects with undecided rate 30%

	$\Delta\nu$	$\Delta\pi$	$\Delta\nu'$	$\Delta\pi'$
Two-stage	2.27	2.27	1.76	1.76
Partially randomized	3.02	NaN	3.29	NaN
Fully randomized	5.51	5.51	3.29	3.29

* the relative preference rate for two treatments remain
  unchanged (47.4% vs. 52.6%), and the base cases are
  the estimated first contrasts and second contrasts of the
  two-stage design in the Cooper's example

Table 4.7: Relative efficiencies of the estimated effects with undecided rate 50%

	$\Delta\nu$	$\Delta\pi$	$\Delta\nu'$	$\Delta\pi'$
Two-stage	1.62	1.62	1.48	1.48
Partially randomized	3.29	NaN	3.59	NaN
Fully randomized	4.66	4.66	3.59	3.59

* the relative preference rate for two treatments remain
  unchanged (47.4% vs. 52.6%), and the base cases are
  the estimated first contrasts and second contrasts of the
  two-stage design in the Cooper's example

Taking the estimated standard errors of estimable effects in the two-stage design as the base case, Table 4.4 shows clearly the relative efficiencies of various estimated effects for the partially randomized design and fully randomized design were more than three times greater than those for the two-stage design.

As shown in Table 4.5, precisions of the estimated effects for the two-stage design show a significant improvement if the proportion of undecided participants could have been reduced to 10% from the original 69.2%, with the relative preference rates for two treatments unchanged (47.4% vs. 52.6%). For instance, relative efficiency of the first contrasts is approximately three times that for the base case and efficiency of the second contrasts increases by 45% of the original one. However, we can see that relative precisions of first selection contrast and second selection contrast for

the partially randomized design become worse if the undecided rate could have been reduced to 10%. Regarding the fully randomized design, Table 4.5 shows a substantial improvement in efficiency of the first contrasts, but a decrease in precision of the second contrasts. Specifically, relative efficiency of the first contrasts is nearly three times that of the base case, while that of the second contrasts decrease by 35% of to base case.

If the undecided rate could have been reduced to 30%, (Table 4.6) precisions of various effects for the two-stage design are enhanced substantially while the benefit to the partially randomized is trivial. Compared to scenario with undecided rate 10%, precision of second contrasts for the two-stage design also increases slightly. For the fully randomized design, relative efficiency of first contrasts increases by 51% of the base case and there is a little difference in the second contrasts.

Table 4.7 presents the relative efficiencies of various effects with 50% of undecided participants. Relative efficiencies of first contrasts and second contrasts for the two-stage design increase by 62% and 48%, respectively. Precisions of estimated effects for the partially randomized design and fully randomized design improve slightly compared to the base case. But the estimated effects in the partially randomized design show the highest precisions if the undecided rate could have been 50%.

In summary, precisions of first contrasts for the two-stage design show a significant improvement if the proportion of undecided participants could have been reduced to

10%, For the partially randomized design, estimable effects show outstanding precisions with 50% undecided participants. Relative efficiencies of the various estimable effects for the partially randomized design is higher than that for the two-stage design if undecided rate could have been not far away from 50%. The fully randomized design has good performance (lower standard error) relative to other designs in all scenarios but may be less acceptable ethically, because of disrespecting the stated preferences of some patients.

Chapter 5

Summary

We have shown earlier the relative efficiencies of the estimable effects with different proportions of undecideds for each design, under equal allocation and unequal allocation. Comparing within each design reveals patterns of the efficiency when the relative preferences and proportion of undecideds vary. Similarly, we also analyzed the relative efficiencies between designs. In practice, one may prefer one design to another depending on the effects which investigator is most interested in or depending on the various circumstances (such as different proportions of undecideds).

Treatment effect measures the effect caused by the treatment itself. When TE is of greatest interest, as we have shown earlier, the conventional design and fully randomized design have the same and higher efficiencies than the other designs. Relative efficiency in the two-stage design is higher than that in the partially randomized design when the proportion of undecideds is less than 50% and vice versa. However, when considering the choice among difference designs, we should not only consider the precisions of effects of interest, other aspects (such as the randomization issue)

should also be taken into consideration. In the conventional design, randomizing the patients without identifying the preference of patients assures the unbiasedness of estimator. However, the patients, who may have a strong preference for one treatment, may refuse to be randomized, which may affect generalizability of results. Same problem exists in the fully randomized design. Although the preferences are identified at baseline in the fully randomized design, patients are all randomized to the treatment. We can also notice that when the undecided participants represent more than half of the participants, efficiency of the TE in the partially randomized design tends to be close to these two designs. Compared with the conventional and fully randomized designs, the partially randomized design allows all patients with preferences to receive their favored treatment, which increases the ethical credibility. Thus, one can consider the partially randomized design when there is a high proportion of undecideds. Otherwise, if the treatment effect is not the only effect which investigators are interested in, then the fully randomized design is a better choice for estimating the TE.

As we shown earlier, variances of the first selection and preference contrasts are same within the design if the effects are estimable. Thus, patterns of relative efficiencies for these two effects are similar except that the PE is not estimable in the partially randomized design. Relative efficiencies are improved as the proportion of undecideds decrease in the two-stage and fully randomized design while relative efficiency in the partially randomized design is not strictly monotonic. The estimate of selection effect in the partially randomized design is partly based on the group of undecideds. Therefore, no information is available when there are no undecided participants for the partially randomized design. Besides, first contrasts could not be

estimated if there only exist undecided participants as preference and selection effect measure the difference in the outcomes between participants prefer treatment A and those prefer treatment B. Thus, the loss of participants with a definite preference would lead to missing values of first contrasts. The efficiencies of selection effect for the partially randomized design and fully randomized design tend to be close to each other when there is a high proportion of undecideds. This is because that when it is the case, both designs include a great portion of undecided participants who are then randomized to the treatment. When the proportion of undecideds is not far from 50%, relative efficiency of the selection effect in the partially randomized design is not far below that in the fully randomized design, so the partially randomized design may be a good choice for those interested in the selection effect. Investigators who are interested in the preference effect only have two choices available, which are two-stage design and fully randomized design. Specifically, when the relative preference is 9:1, relative efficiency of the first contrasts for the two-stage design is very low (below 0.05), which is obviously not reasonable to be considered. Relative efficiency for the fully randomized design is generally above one. Instead, relative efficiency for the two-stage design is greater than 1 only when there are less than 50% of undecideds and relative preference rate is between 30% and 70%. Therefore, one may consider the two-stage design when the fraction of undecideds is relatively small and no substantial difference in the preference pattern among participants, and it is relatively more difficult to implement the fully randomized design.

The second selection and preference contrasts are concerned with the differences

67

in outcomes between participants with a definite preference and those with no preference, which turns out that they are only able to be estimated if both types of participants are present. Regarding the second selection contrast, relative efficiencies for the partially randomized design and the fully randomized design are same and substantially higher than relative efficiency in the two-stage design. Thus, investigators who are interested in the second selection effect may consider partially or fully randomized design. We can also notice that relative efficiency of the second contrasts for the two-stage design is above 1 only when the proportion of undecided participants is slightly less than 50% and the relative preference is around 50:50. Thus, if the second preference effect is of primary interest, two-stage design can also be considered when it is the case.

Comparing between the unequal and equal randomization for the two-stage design, relative efficiencies of the first and second contrasts in the unequal randomization are close to those in the equal case, with greater gains achieved when participants are far from equally preferred. One may consider assigning more undecideds to the treatment which is more preferred among participants with a definite preference. On the other hand, relative efficiencies of the first and second contrasts in the partially and fully randomized designs become worse when participants are not equally randomized.

Comparisons within each design in the case of unequal randomization lead to similar results to the case of equal randomization. One difference is that relative efficiencies with respect to second selection contrast are not same in the partially and fully randomized designs. However, the difference is subtle. Thus, choice between

those two designs depends on other factors. Neither designs assure an unbiased estimation of selection effect. The estimations are both potentially biased, as noted in Chapter 2. Fully randomized design involves randomizing patients who have a strong preference, which leads to ethics problems. However, the partially randomized design respects the preferences of all patients and only randomize the patients without a treatment preference. One may consider choosing the partially randomized design when facing the difficulties of allocating the patients with a strong preference.

Patient preferences in the random arm are explicitly identified in the choice arm of the two-stage randomized design, and preferences are all identified in the partially randomized design and fully-randomized design. In contrast, the information with respect to the patient preferences can be indirectly revealed by the observed treatment patterns in the conventional trials. To measure the treatment effect for the conventional trials, three traditional analyses are available, which are intention-to-treat (ITT) analysis, per-protocol (PP) analysis and as treatment (AT) analysis. Consider two randomization groups, with one group assigned treatment T0 and the other T1. ITT measures the treatment effect with respect to the assigned treatment, so this analysis ignores the information about the actual treatment patients received. PP method only concerns the patients who comply with their assigned treatment. As for the AT analysis, treatment effect is measured from a respective of received treatment. To measure the treatment effect involving these scenarios, the latent class instrumental variable (IV) method was formulated by Baker *et al.* (2016) . Four latent classes were considered in this method, which were (T0, T0), (T0,T1), (T1,T0), (T1,T1) and are named as "never-taker", "complier", "defier" and "always-taker",

respectively. (Angrist *et al.*, 1996) For instance, compliers receive treatment T1 if assigned to treatment T0 while defiers receive T0 when assigned to T1. Considering the outcomes as binary, the corresponding intent-to-treat, per-protocol, and as-treated measures are formulated.

A "preference-based analysis" for conventional trials, which was proposed by Walter *et al.* (2006) used the observed treatment patterns to estimate treatment effects in conventional trials. In this paper, five preferences groups are considered, which are compliers, preferers (A or B) and insisters (A or B). Compliers accept either treatment assigned to them; Preferers accept their desired treatment or otherwise refuse; Insisters would accept the treatment they prefer, if it is not offered, they insist on their preferences. Compared with ITT, AT, PP, preference-based analysis focuses on estimating the treatment effect among the "compliers", who are committed to the treatment prescribed. In addition, this method also provides an estimation of the preference distribution among these five preferences groups.

One limitation of this thesis is that only participants who agreed to be in a trial are discussed. This could lead to a lack of generalizability. A comprehensive cohort design could improve generalizability by including participants who refused to be in a trial. At the first stage of the comprehensive cohort design, all eligible participants are split into two groups in a non-random way. Participants who refuse to be randomized are still kept in the study and they can choose the treatment based on the preference. Those who consent are randomized to the two treatments. Similar to the two-stage design, the comprehensive cohort design comprises a randomized group and

a "choice" group. However, the "choice" group in the comprehensive cohort design involves selection bias into the study due to the non-random way of splitting the participants. Therefore, the preference distribution observed in the non-randomized group cannot be used to estimate the preference effect. Specifically, a difference might exist in the preference between randomized participants and those not randomized .

There are a lot of issues that may cause the difference. Rothwell (2005) classified these issues into six categories, which are setting of the trial,"selection of patients", "characteristics of randomized patients", "differences between the trial protocol and routine practice","outcome measures and follow-up", "adverse effects of treatment". Regarding the "setting of a trial", differences between health-care systems, national differences in the method of diagnosis and management, and selection of centers and clinicians could all possibly have an impact on the external validity (generalizability). In addition to the highly selective eligibility criteria, exclusion of patients before eligibility may also undermine external validity. Specifically, Rothwell (2005) showed that the proportion of a particular disorder in the local community served by a participating center who are considered for recruitment into a trial will often be below 1% of the all available patients. As a result of eligibility criteria, patients in trials are usually healthier, younger, and of higher social status than those not in trials. Mant (1999) pointed out that even if it is the case, it does not matter when the conditions are well defined and interventions to be assessed are simple. However, if we take the heterogeneity of effect for individual patients into account, then it still seems likely to make a difference in outcome between the patients in trials and those not.

Therefore, if patients in trials are not representative, then estimated preference rates will be biased by selection bias. Also, even if no significant difference in preferences exists, the treatment outcomes might be different in these two groups (in trial, vs. not in trial). An extreme case could be that if the participants in a trial show a better outcome in one treatment, but the other treatment seems more helpful for those not in a trial. Therefore, in addition to the impact on the magnitude of effects, the direction of effects might also be changed. Overall, clinicians may take into account of the external validity before applying the treatment outcomes obtained in trials into patients.

Bibliography

Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of casual effects using instrumental variables. *Journal of the American Statistical Association*, **91**, 444–455.

Baker, S. G., Kramer, B. S., and Lindeman, K. S. (2016). Latent class instrumental variables: a clinical and biostatistical perspective. *Statistics in Medicine*, **35**, 146–160.

Brewin, C. R. and Bradley, C. (1989). Patient preferences and randomised clinical trials. *BMJ: British Medical Journal*, **299**, 313–315.

Cameron, B., Peduzzi, P., and Esserman, D. (2018). Extensions to the two-stage randomized trial design for testing treatment, self-selection, and treatment preference effects to binary outcomes. *Statistics in Medicine*, **37**, 3147–3178.

Cooper, K. G., Grant, A. M., and Garratt, A. M. (1997). The impact of using a partially randomized patient preference design when evaluating alternative managements for heavy menstrual bleeding. *BJOG: An International Journal of Obstetrics & Gynaecology*, **104**, 1367–1372.

Cooper, K. G., Parkin, D. E., Garratt, A. M., and Grant, A. M. (1999). Two-year

follow up of women randomized to medical management or transcervical resection of the endometrium for heavy menstrual loss: clinical and quality of life outcomes. *BJOG: An International Journal of Obstetrics & Gynaecology*, **106**, 258–265.

Mant, D. (1999). Can randomized trails inform clinical decisions about individual patient? . *The Lancet*, **353**, 743–746.

Rothwell, P. (2005). External validity of randomized controlled trials:"To whom do the results of this trial apply?" . *Lancet*, **365**, 82–93.

Rücker, G. (1989). A two-stage trial design for testing treatment, self-selection and treatment preference effects. *Statistics in medicine*, **8**, 477–485.

Torgerson, D. J., Klaber-Moffett, J., and Russell, I. T. (1996). Patient preferences in randomised trials: threat or opportunity? . *Journal of Health Services Research & Policy*, **1**, 194–197.

Walter, S., Guyatt, G., Montori, V. M., Cook, R., and Prasad, K. (2006). A new preference-based analysis for randomized trials can estimate treatment acceptability and effect in compliant patients. *Journal of Clinical Epidemiology*, **59**, 685–696.

Walter, S., Turner, R., Macaskill, P., McCaffery, K., and Irwig, L. (2017a). Beyond the treatment effect: evaluating the effects of patient preferences in randomized trials. *Statistical Methods in Medical Research*, **26**, 489–507.

Walter, S. D., Turner, R. M., Macaskill, P., McCaffery, K. J., and Irwig, L. (2017b). Estimation of treatment preference effects in clinical trials when some participants are indifferent to treatment choice. In *BMC medical research methodology*.